



UNIVERSITY OF
LIVERPOOL

AN INVESTIGATION OF GIARDIASIS AND CRYPTOSPORIDIOSIS IN MALAWI AND CAMBODIA

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of
Doctor in Philosophy

Chadaporn Nuchjangreed

March 2018

AUTHOR'S DECLARATION

Apart from the help and advice acknowledged, this thesis represents the unaided
work of the author

.....

Chadaporn Nuchjangreed
March 2018

This research was carried out in the Department of Infection Biology,
University of Liverpool

ACKNOWLEDGEMENTS

This study was carried out at the department of Infection Biology, Institute of Infection and Global Health, University of Liverpool, UK. Otherwise, it was financially supported by a grant from the Royal Thai Government, Thailand. I would like to express my sincere thanks to the organizations above, and to all who have contributed to my work and life here in the UK.

Thank you to a huge number of people who have supported, encouraged and I am very sure despaired that I would never complete this document.

I greatly appreciate the help of the following people:

First of all, I would like to thank Professor Jonathan M. Wastling, my main supervisor, who has given me the benefit of all his specialist experience in my field of study, his support and useful ideas regarding my research. His kindness and hospitality was tremendously appreciated.

So many thanks to Professor Nigel Cunliffe, my second supervisor, for his support, allowing me the use of the human samples used in this research and his useful ideas regarding my thesis. Moreover, his advice, his help with the corrections, his patience and time in reading and comments throughout my thesis. Without his helpful advice and generous consideration, my thesis would never have been completed. Also, Dr. Nick Beeching for his support and useful ideas regarding my research. Winifred Dove, one of the team in this *Cryptosporidium* project for preparing all the samples, her advice, support and helped me with statistical analysis. Dr. Wendi Bailey for training in the direct immunofluorescence assay (DFA) and providing me with laboratory facilities in the Diagnostic Laboratory, Liverpool School of Tropical Medicine, University of Liverpool.

A special thank goes to Professor Chris Probert, one of the team in this *Cryptosporidium* project for introducing me to a novel method for rapidly diagnosing the causes of diarrhoea by using Gas Chromatography Mass Spectrophotometry (GCMS) and providing me with his laboratory facilities in the department of Gastroenterology, The Henry Wellcome Laboratory, University of Liverpool and also for his precious advice and comments on methods, statistics and data analysis in Chapter 5.

A huge thank you to Dr. Nadine Randle, for her helpful and supportive advice in the molecular laboratory. When my laboratory work was in trouble, she was always there with ideas and helpful suggestions. I am also thankful for her patience and time in correcting my thesis, invaluable scientific input and comments throughout the entire project.

Many thanks also to Dr. Andrew Jackson and Dr. Simon Clegg for their help about phylogenetic analysis. Thanks to Dr. Stu Armstrong and Dr. Dong Xia for their precious comments and help during the project. Thanks to Nicholas Brown for his patience and time in correcting my thesis. Thanks to Kittichai Unjit for listening to me when I had problems in academic and private issues and also always helpful and supportive of me.

I would like to thank all staff and Ph.D students in the Infection Biology group that helped and supported me through my Ph.D experience. Thanks to all my friends and people for their valuable support and kind suggestions during my time in the Veterinary School, University of Liverpool.

A huge thank you to Dr. Neil Blake for all his support, knowledge and kind help. A special thanks to Professor Rob Christley, the internal examiner and Professor Glenn McConkey, the external examiner. Without their helpful advice and generous consideration, my thesis would never have been completed. Their kindness and hospitality was tremendously appreciated.

Last but not least, I would like to thank Mr. Adam Gordon for his help with the corrections, his patience and time in proof reading and comments throughout my thesis. I really appreciate his encouragement, great support and positivity. I am also grateful and thankful to my family, especially my mom for her love, invaluable support and everything during my time here in the UK, without everyone I mentioned above this project wouldn't have been possible

ABSTRACT

The parasites *Giardia duodenalis* and *Cryptosporidium* spp., which infect a broad range of vertebrates, are major causes of diarrhoeal disease in humans worldwide. The aim of this thesis was to understand the zoonotic transmission potential of Giardiasis and Cryptosporidiosis in children in Malawi and Cambodia using molecular epidemiological tools. The pattern of faecal Volatile Organic Compounds (VOCs) from children with Cryptosporidiosis in Malawi was also examined.

In Chapter 2, faecal samples collected from children under 5 years of age with diarrhoea, living in diverse geographical regions in Malawi, were screened for both parasites microscopically using the Direct Fluorescence Assay (DFA). Nested PCR was used for subtyping *Giardia* (the *bg*, *tpi* and *gdh* genes) and *Cryptosporidium* spp. (SSU-rRNA and GP60 genes). The prevalence of *G. duodenalis* (11.4%, 15/132) was significantly lower than the prevalence of *Cryptosporidium* spp. (23.5%, 31/132) by DFA. In contrast, the prevalence of *G. duodenalis* (28%, 56/200) was significantly higher than the prevalence of *Cryptosporidium* spp. (11%, 22/200) by nested PCR. Mixed infections with both *Giardia* assemblage A and B were predominant. The predominance of *C. hominis* indicated that the anthroponotic route plays a major role of *Cryptosporidium* transmission in Malawi.

In Chapter 3, the prevalence and identity of *G. duodenalis* assemblages among isolates from children in Cambodia was investigated. Both assemblages A and B were common in Cambodia, however, assemblage B was the most predominant. These data indicate that anthroponotic routes play a major role in *Giardia* transmission in Cambodia. Nevertheless, *Giardia* parasites of assemblage B demonstrated higher genetic variation than the other assemblages. Because the major transmission route of *Giardia* is likely anthroponotic, public health policies should focus on improvements in sanitation and hygiene rather than changes in animal/meat processing.

In Chapter 4, I determined genetic variation among *Giardia* isolates from Malawi and Cambodia. A multiple alignment was performed and phylogenetic trees were generated. The data demonstrated that *Giardia* assemblage A and B were prevalent in children in both countries, with a predominance of assemblage B. These findings suggest that anthroponotic transmission could be a dominant transmission route for giardiasis in both countries.

In Chapter 5, I studied stool samples from patients with confirmed *Cryptosporidium* infection and negative controls. I tested the hypothesis that the pattern of excreted VOCs could be specific to certain types of diarrhoeal infection. The abundance of several VOCs were significantly different in the two groups; cyclopentane and 3-hydroxy-2-butanone were associated with *Cryptosporidium* positive samples, and 1-propanol with negative control samples. The presence of cyclopentane and 3-hydroxy-2-butanone and the absence of 1-propanol in *Cryptosporidium* positive faecal samples could form the basis of a diagnostic test.

In Conclusion, my findings suggest that *G. duodenalis* assemblage B was the predominant assemblage in Malawi and Cambodia. *C. hominis* was the predominant species in Malawi. Both parasites likely undergo predominantly anthroponotic transmission in these populations. Public health measures targeted at hand washing, improving sanitation, and providing clean drinking water are important strategies for the control of Giardiasis and Cryptosporidiosis infection in these countries.

TABLE OF CONTENTS

AUTHORS DECLARATION.....	ii
ACKNOWLEDGEMENTS.....	iii
ABSTRACT.....	v
TABLE OF CONTENTS	vi
LIST OF FIGURES.....	xv
LIST OF TABLES.....	xix
LIST OF ABBREVIATIONS.....	xxii

CHAPTER ONE: GENERAL INTRODUCTION

1.1 The global importance of intestinal parasites	1
1.2 Introduction to <i>Giardia duodenalis</i> and Giardiasis	2
1.2.1 Biology and the life cycle of <i>Giardia</i>	2
1.2.2 Transmission of <i>Giardia</i>	4
1.2.3 Taxonomy and genotyping of <i>Giardia</i>	5
1.2.4 Zoonotic Transmission	6
1.2.5 Symptoms and clinical features of Giardiasis	6
1.2.6 Pathogenesis of Giardiasis	8
1.2.7 Treatment of Giardiasis	8
1.2.8 Prevention of Giardiasis	9
1.3 Introduction to <i>Cryptosporidium</i> and Cryptosporidiosis	10
1.3.1 Biology and life cycle of <i>Cryptosporidium</i> spp.	11
1.3.2 Transmission of <i>Cryptosporidium</i> spp.	12

1.3.3 Taxonomy and Genotyping of <i>Cryptosporidium</i> spp.	14
1.3.4 Zoonotic transmission	14
1.3.5 Symptoms and clinical features of Cryptosporidiosis	15
1.3.6 Pathogenesis of Cryptosporidiosis	16
1.3.7 Treatment of <i>Cryptosporidium</i>	16
1.3.8 Prevention of <i>Cryptosporidium</i>	18
1.3.9 <i>Cryptosporidium</i> species/genotypes in humans	19
1.3.10 <i>Cryptosporidium hominis</i> infections	21
1.3.11 Human infections with other <i>Cryptosporidium</i> spp.	23
1.3.12 <i>Cryptosporidium</i> species and genotypes in humans in developing countries	24
1.3.13 Zoonotic Cryptosporidiosis in developing countries	25
1.3.14 Molecular epidemiology and public health implications	26
1.3.14.1 Small subunit (SSU) rRNA	26
1.3.14.2 Oocyst wall protein (COWP) gene	26
1.3.14.3 60 kDa glycoprotein (GP60 or GP40/15)	27
1.3.14.4 The identification of microsatellite and minisatellite sequences	27
1.3.14.5 Multilocus sequence typing (MLST)	28
1.4 Detection of <i>Giardia duodenalis</i> and <i>Cryptosporidium</i> spp. in clinical samples	29
1.4.1 Microscopy	29
1.4.2 Immunological methods	29
1.4.3 Molecular methods	30

1.5 Current status of <i>Giardia duodenalis</i> and <i>Cryptosporidium</i> spp. infection in Africa	30
1.5.1 <i>Giardia duodenalis</i>	30
1.5.2 <i>Cryptosporidium</i> spp.	33
1.6 Current status of <i>Giardia duodenalis</i> and <i>Cryptosporidium</i> spp. infection in South East Asia	39
1.6.1 <i>Giardia duodenalis</i>	39
1.6.2 <i>Cryptosporidium</i> spp.	44
1.7 Development of metabolomics method for the analysis of human faecal samples using <i>Cryptosporidium</i> -infected samples	47
1.8 Aims and objectives	50
 CHAPTER TWO: GIARDIASIS AND CRYPTOSPORIDIOSIS IN MALAWI	
2.1 INTRODUCTION	51
2.1.1 Diarrhoeal diseases in Africa	51
2.1.2 <i>Giardia duodenalis</i> and <i>Cryptosporidium</i> spp. infection in Africa	51
2.2 AIM OF STUDY	54
2.3 MATERIALS AND METHODS	55
2.3.1 Source of samples	55
2.3.2 Direct Fluorescence Assay	56
2.3.3 DNA extraction from faecal samples	57
2.3.4 Multilocus genotyping of <i>Giardia</i> and <i>Cryptosporidium</i>	58
2.3.4.1 Genotyping and subtyping of <i>Giardia</i>	58
2.3.4.2 PCR amplification of small subunit-rDNA gene (18S SSU-rDNA)	59

2.3.4.3 Triose phosphate isomerase (<i>tpi</i>) gene	60
2.3.4.3.1 <i>Giardia</i> Assemblage A specific PCR	60
2.3.4.3.2 <i>Giardia</i> Assemblage B specific PCR	60
2.3.4.4 β -giardin (<i>bg</i>) gene	61
2.3.4.5 Glutamate dehydrogenase (<i>gdh</i>) gene	61
2.3.5 Genotyping and subtyping of <i>Cryptosporidium</i> spp.	62
2.3.5.1 Primers for genotyping and subtyping of <i>Cryptosporidium</i> spp.	62
2.3.5.2 SSU-rRNA gene	63
2.3.5.3 Restriction Fragment Length Polymorphism (RFLP) of the SSU PCR	63
2.3.5.4 GP60	64
2.3.6 Purification of PCR products	65
2.3.7 Sequencing of PCR products and DNA Sequence analysis	65
2.3.8 Statistical analyses	65
2.3.9 Ethical approval for the studies	66
2.4 RESULTS	67
2.4.1 Direct Fluorescence Assay for <i>G. duodenalis</i> and <i>Cryptosporidium</i> spp.	67
2.4.2 Multilocus genotyping and subtyping of <i>G. duodenalis</i>	69
2.4.2.1 Analysis of the small subunit-rDNA gene (18S SSU -rRNA)	69
2.4.2.2 Analysis of the Triose Phosphate Isomerase gene (<i>tpi</i> gene)	69
2.4.2.3 Analysis of the β -giardin gene	71
2.4.2.4 Analysis of glutamate dehydrogenase (<i>gdh</i> gene)	72

2.4.3 Genotyping and subtyping of <i>Cryptosporidium</i> spp.	73
2.4.3.1 PCR amplification of SSU-rRNA gene	73
2.4.3.2 PCR- Restriction Fragment Length Polymorphism	73
2.4.3.3 PCR amplification of GP60	76
2.4.3.4 SSU rRNA-based PCR-RFLP, GP60 and DNA sequence analysi	76
2.4.4 The correlation between gender, rotavirus co-infection, age and <i>G. duodenalis</i> assemblages/ <i>Cryptosporidium</i> spp.	77
2.4.4.1 The correlation between gender and <i>G. duodenalis</i> assemblages/ <i>Cryptosporidium</i> spp.	77
2.4.4.2 The correlation between rotavirus co-infection and <i>G. duodenalis</i> assemblages/ <i>Cryptosporidium</i> spp.	77
2.4.4.3 The correlation between age and <i>G. duodenalis</i> assemblages/ <i>Cryptosporidium</i> spp.	78
2.4.5 Age distribution of <i>G. duodenalis</i> and <i>Cryptosporidium</i> spp. cases in Malawian Children	83
2.4.6 Seasonal distribution of <i>G. duodenalis</i> and <i>Cryptosporidium</i> spp. in Malawian Children	84
2.5 DISCUSSION	85
CHAPTER THREE: MULTILOCUS GENOTYPING OF <i>GIARDIA DUODENALIS</i> AMONG CHILDREN IN CAMBODIA	
3.1 INTRODUCTION	96
3.1.1 Intestinal parasite infections in Cambodia	96

3.1.2 Molecular genotyping for <i>Giardia</i> in Cambodia and countries in Southeast Asia	97
3.1.3 Using a multi-locus sequence typing (MLST) for <i>Giardia duodenalis</i> infection in Cambodia and countries in Southeast Asia	98
3.2 AIMS OF THE STUDY	99
3.3 MATERIALS AND METHODS	101
3.3.1 Source of samples	101
3.3.2 DNA samples	102
3.3.3 Clinical information	103
3.3.4 Multilocus genotyping and subtyping of <i>Giardia</i>	103
3.3.5 Triose phosphate isomerase (<i>tpi</i>) gene	103
3.3.5.1 <i>Giardia</i> Assemblage A specific PCR	103
3.3.5.2 <i>Giardia</i> Assemblage B specific PCR	103
3.3.6 β -giardin gene	104
3.3.7 Glutamate dehydrogenase (<i>gdh</i>) gene	104
3.3.8 Purification of PCR products	104
3.3.9 Sequencing of PCR products and DNA Sequence analysis	104
3.3.10 Statistical analyses	104
3.3.11 Ethical consideration for the study	104
3.4 RESULTS	105
3.4.1 Overall prevalence of <i>G. duodenalis</i> assemblages	105
3.4.2 <i>G. duodenalis</i> sub-assemblages and genotypes by locus	107
3.4.2.1 Triose phosphate isomerase (<i>tpi</i> gene)	107
3.4.2.2 <i>Beta-giardin</i> gene	108

3.4.2.3 <i>Glutamate dehydrogenase</i> gene	109
3.4.3 The associations between patient characteristics and <i>G. duodenalis</i> infection in children	109
3.4.4 The association between the duration of diarrhoea and infection <i>G. duodenalis</i>	113
3.4.5 The association between the duration of abdominal pain and infection <i>G. duodenalis</i>	114
3.4.6 The association between the different geographical locations and infection with <i>G. duodenalis</i> (<i>tpi</i> gene)	115
3.5 DISCUSSION	117
 CHAPTER FOUR: COMPARISON OF EVOLUTIONARY ORIGIN OF <i>GIARDIA</i> FROM MALAWI AND CAMBODIA	
4.1 INTRODUCTION	123
4.2 AIMS OF THE STUDY	124
4.3 MATERIALS AND METHODS	125
4.3.1 Isolates	125
4.3.2 Purification of the secondary PCR product	125
4.3.3 Preparation of DNA sequences	125
4.3.4 Blast and alignment	125
4.3.5 Evolutionary informatics	126
4.3.6 Phylogenetic tree construction	127
4.4 RESULTS	130
4.4.1 <i>G. duodenalis</i> sub-assemblages and genotypes by locus	130
4.4.1.1 <i>Triose phosphate isomerase</i> gene	130

4.4.1.2 <i>Beta-giardin</i> gene	134
4.4.1.3 <i>Glutamate dehydrogenase</i> gene	137
4.5 DISCUSSION	140
 CHAPTER FIVE: AN INVESTIGATION OF THE USE OF GAS CHROMATOGRAPHY MASS SPECTROPHOTOMETRY (GCMS) FOR DIAGNOSING <i>CRYPTOSPORIDIUM</i> INFECTION IN CHILDREN IN MALAWI	
5.1 INTRODUCTION	148
5.2 AIMS AND OBJECTIVES	151
5.3 MATERIALS AND METHODS	151
5.3.1 Sample Preparation	151
5.3.2 VOCs Extraction	153
5.3.3 Analysis of volatile organic compounds	153
5.3.4 AMDIS	154
5.3.5 MetaboAnalyst 2.0 programme	155
5.3.6 Partial least squares Discriminant analysis (PLS-DA)	156
5.3.7 Statistical analysis	157
5.4 RESULTS	158
5.4.1 Non infected <i>Cryptosporidium</i> subjects (negative control group)	158
5.4.2 Specific changes in VOCs in <i>Cryptosporidium</i> positive samples	158
5.4.3 Statistical analysis of <i>Cryptosporidium</i> samples and negative controls	159
5.4.3.1 Analysis by Chi square test	159
5.4.3.2 Partial Least Squares-Discriminant Analysis (PLS-DA)	159

5.5 DISCUSSION	163
CHAPTER SIX: DISCUSSION AND CONCLUSIONS	168
CONCLUSIONS	181
APPENDIX	184
Appendix 1 Demographic data of symptomatic children from Malawi in relation to <i>Giardia</i> and <i>Cryptosporidium</i> positive by PCR and DFA	185
Appendix 2 The distribution of <i>Giardia</i> assemblage A and B from Malawi and Cambodia based on the different loci and mixed infection	188
Appendix 3 Sequence chromatograms of nucleotide variations. Chromatogram of sequences generated from <i>Giardia</i> Assemblage A and B using multi-locus markers	201
Appendix 4 Multiple alignments of gene (<i>tpi</i> , <i>bg</i> and <i>gdh</i>) sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B	212
Appendix 5 Total number of compounds occurring in each <i>Cryptosporidium</i> patient samples and non-infected samples	229
Appendix 6 A selection of certificates from some of my presentations, training courses and workshops	234
REFERENCES	241

LIST OF FIGURES

- Figure 1.1** The life cycle of *G. duodenalis*. The cysts when passed from faeces, can survive in the environment for many months (1). The cysts from contaminated water or food are ingested by the host and then excysted in the duodenum (2). Trophozoites are released from each cyst (3). The trophozoites multiply asexually by longitudinal binary fission (4). The multiplication continues and the cysts are passed with stools, and can then be transmitted to new hosts (5). Picture taken from <http://www.dpd.cdc.gov/dpdx/giardiasis> 4
- Figure 1.2** The Life cycle of *Cryptosporidium* spp. Picture taken from <http://www.dpd.cdc.gov/dpdx/cryptosporidiosis> 12
- Figure 2.1** Map of Malawi showing the regions of Karonga (1), Lilongwe (2) and Mangochi (3) taken from https://commons.wikimedia.org/wiki/file:Malawi_map.png 56
- Figure 2.2** *G. duodenalis* cysts by DFA: The *G. duodenalis* cysts are oval in shape with a bright apple green colour approximately 8-12 µm long (panel A). *Cryptosporidium* oocysts are round to slightly oval in shape with a bright apple green colour, approximately 2-6 µm in diameter (Panel B) negative result (neither *G. duodenalis* nor *Cryptosporidium* spp. (C) 68
- Figure 2.3** Visualisation of PCR-RFLP of *Cryptosporidium* spp. based on SSU-rRNA gene from Lilongwe samples, Lane L: Ladder (100 bp marker), Lane 2, 4, 6, 8, and 9 = *C. hominis*; Lane 1 and 7: Mixed infection with both genotype, Lane 3 and 5 = *C. parvum*, Lane 10 = +ve (*C. hominis*), Lane 11 = +ve (*C. parvum*), Lane 12 = -ve, Notice: *SspI* (upper gel), *VspI* (lower gel). 75
- Figure 2.4** The number of positive samples for *G. duodenalis* (A) and *Cryptosporidium* (B) in male and female children 79
- Figure 2.5** The number of positive samples for *G. duodenalis* (A) and *Cryptosporidium* (B) found in two groups of children with or without rotavirus infection 80

- Figure 2.6** The number of positive samples for *G. duodenalis* (A) and *Cryptosporidium* (B) found in different age groups. 81
- Figure 2.7** The percentage of *G. duodenalis* assemblages/ found in male and female patients (A), the percentage of *G. duodenalis* assemblages found in two groups of children with or without rotavirus infection (B), and the percentage of *G. duodenalis* assemblages found in different age groups (C) 82
- Figure 2.8** Age distribution of *G. duodenalis* (n=56) and *Cryptosporidium* spp. (n=32) cases in Malawian Children 83
- Figure 2.9** Seasonal distribution by month of *G. duodenalis* (n=56) and *Cryptosporidium* spp. (n=32) cases in Malawian Children 84
- Figure 3.1** Map of Cambodia showing the regions of Banteay Meanchey (1), Battambang (2) Siem Reap (3), Kampong Thom (4) and Kompong Cham (5) adapted from <https://heifer12x12.com/map-of-cambodia/> 102
- Figure 3.2** Overall sub-assemblage typing results at the *tpi* locus in 12 assemblage A (A) and 28 assemblage B isolates (B). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity 107
- Figure 3.3** Overall sub-assemblage typing results at the *bg* locus in 4 assemblage A (A) and 15 assemblage B isolates (B). Isolates (with number and percentage reported) grouped by sub-assemblage similarity 108
- Figure 3.4** Overall sub-assemblage typing results at the *gdh* locus in 7 assemblage A (A) and 38 assemblage B isolates (B). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity 109
- Figure 4.1** Phylogenetic relationships of *G. duodenalis* isolates showing the maximum likelihood tree of the *tpi* gene sequence of 89 isolates (Malawi =49 samples and Cambodia =40 samples). The numbers on branches are bootstrap values and the scale bar shows an evolutionary distance of 0.10 nucleotides per position in the sequence. Number above or below branches

show bootstrap value (%). One representative sequence from each identified genotype is demonstrated with the total number of isolates sharing the same sequence reported in parentheses. The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates. The accession number obtained from GenBank are inserted as the reference sequences. Optimal nucleotide substitution model: Tamura-Nei parameter with gamma distribution and invariant sites. Estimated from 1,000 resampling of the sequence data.

133

Figure 4.2 Phylogenetic relationships of *G. duodenalis* isolates showing the maximum likelihood tree of the *bg* gene sequence of 33 isolates (Malawi =14 samples and Cambodia = 19 samples). The numbers on branches are bootstrap values and the scale bar shows an evolutionary distance of 0.10 nucleotides per position in the sequence. Number above or below branches show bootstrap value (%). One representative sequence from each identified genotype is demonstrated with the total number of isolates sharing the same sequence reported in parentheses. The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates. The accession number obtained from GenBank are inserted as the reference sequences. Optimal nucleotide substitution model: Tamura-Nei parameter with gamma distribution and invariant sites. Estimated from 1,000 resampling of the sequence data.

136

Figure 4.3 Phylogenetic relationships of *G. duodenalis* isolates showing the maximum likelihood tree of the *gdh* gene sequence of 83 isolates (Malawi =40 samples and Cambodia =43 samples). The numbers on branches are bootstrap values and the scale bar shows an evolutionary distance of 0.05 nucleotides per position in the sequence. Number above or below branches show bootstrap value (%). One representative sequence from each identified genotype is demonstrated with the total number of isolates sharing the same sequence reported in parentheses. The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates. The accession numbers obtained from GenBank are inserted as the reference sequences. Optimal nucleotide substitution model: Kimura-2 parameter with gamma distribution and invariant sites. Estimated from 1,000 resampling of the sequence data.

139

Figure 5.1 The flowchart of analysis of volatile organic compounds for identification of potential biomarkers of Cryptosporidiosis	156
Figure 5.2 Chromatograms of control sample and <i>Cryptosporidium</i> sample. The figure shows two chromatograms produced by the GCMS analysis. Chromatogram A shows the results of a control sample (negative control, no <i>Crypto</i>) and chromatogram B shows the results for a <i>Cryptosporidium</i> sample (positive control). Both show percentage abundance of ions across retention time. Both having a scanning mass range of 17 to 350 ions. Notice; an internal standard was not used.	160
Figure 5.3 Show the important features identified by PLS-DA. It indicates the relative concentrations of the corresponding metabolite in the compounds found in the <i>Cryptosporidium</i> samples (group 1) and in the compounds found in the controls (group 2)	162
Figure 3A. Chromatogram of a sequence generated from <i>Giardia</i> Assemblage B using <i>tpi</i> gene (forward primer)	201
Figure 3B. Chromatogram of a sequence generated from <i>Giardia</i> Assemblage A and B using <i>bg</i> gene (reward primer)	204
Figure 3C. Chromatogram of a sequence generated from <i>Giardia</i> Assemblage A and B using <i>gdh</i> gene (reward primer)	208
Figure 4A. Multiple alignments of <i>tpi</i> sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B.	212
Figure 4B. Multiple alignments of <i>bg</i> sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B.	217
Figure 4C. Multiple alignments of <i>gdh</i> sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B.	224

LIST OF TABLES

Table 1.1 Occurrence of <i>Cryptosporidium</i> spp. in immunocompetent individuals in developing countries	20
Table 1.2 Occurrence of <i>Cryptosporidium</i> spp. in immunocompromised individuals in developing countries	21
Table 1.3 Summary of studies of <i>Giardia duodenalis</i> infection in Africa	32
Table 1.4 Summary of studies of <i>Cryptosporidium</i> spp. infection in Africa	36
Table 1.5 Current status of <i>Giardia duodenalis</i> infection in Asia	42
Table 1.6 Summary of <i>Cryptosporidium</i> spp. infection in Asia	46
Table 2.1 Lists of primers for genotyping and subtyping of <i>Giardia</i>	59
Table 2.2 Lists of primers for genotyping and subtyping of <i>Cryptosporidium</i> spp.	62
Table 2.3 Patterns of RFLP digestion for <i>Cryptosporidium</i>	64
Table 2.4 The prevalence of Giardiasis and Cryptosporidiosis infections in children with diarrhoea under 5 years of age from three regions in Malawi	68
Table 2.5 The detection of <i>G. duodenalis</i> in children under 5 year of age from three regions in Malawi using SSU-rRNA PCR	69
Table 2.6 <i>G. duodenalis</i> assemblages (<i>tpi</i> gene) in children under 5 year of age from three regions in Malawi	70
Table 2.7 <i>G. duodenalis</i> genotype by β -giardin in children under 5 year of age from three regions in Malawi	71
Table 2.8 <i>G. duodenalis</i> genotyped by <i>gdh</i> gene in children under 5 year of age from three regions in Malawi	72

Table 2.9 The detection of <i>Cryptosporidium</i> spp. (SSU-rRNA and GP60 genes) in children under 5 year of age from three regions in Malawi	73
Table 2.10 Genotyping results by using SSU rRNA-based PCR-RFLP, GP60 and sequencing for <i>Cryptosporidium</i> spp. from 3 regions in Malawi	76
Table 3.1 The prevalence of <i>G. duodenalis</i> assemblages examined in 484 faecal samples by nested PCR	106
Table 3.2 Assemblage typing results by three different markers in the four samples with a mixed assemblage A and B infection	106
Table 3.3 Characteristic of children infected with <i>Giardia</i> (<i>tpi</i> gene) (n=74)	110
Table 3.4 Associations between patient characteristics and <i>G. duodenalis</i> , sub-assemblage infection based on <i>gdh</i> gene from children aged less than 16 years (n=74)	112
Table 3.5 Frequency of samples infected with <i>G. duodenalis</i> (<i>tpi</i> gene) on diarrhoea duration days specified by sub-assemblages (AI, AII, BIV H43 (EF507682), B (VAN/94/UBC/122), B (VAN/94/UBC/122), BIV (KC96064.1) and B (un-identified)	113
Table 3.6 Frequency of samples infected with <i>G. duodenalis</i> (<i>tpi</i> gene) on abdominal pain duration days specified by sub-assemblages (AI, AII, BIV H43(EF507682), B(VAN/94/UBC/122), B (VAN/94/UBC/122), BIV (KC96064.1) and B (un-identified)	115
Table 3.7 The association between the different geographical locations and infection with <i>G. duodenalis</i> (<i>tpi</i> gene) by Assemblages A, B and AB mixed (n = 74)	116
Table 4.1 Details of DNA sequences from GenBank	128
Table 4.2 Summary of the 89 isolates successfully sequenced at the <i>tpi</i> gene	131
Table 4.3 Summary of the 33 isolates successfully sequenced at the <i>bg</i> gene	135
Table 4.4 Summary of the 83 isolates successfully sequenced at the <i>gdh</i> gene	138

Table 5.1 Demographics of the study participants	152
Table 5.2 Biologically volatile organic compounds identified using AMDIS. Twenty six biologically compounds determined by Chi-Square test, these compounds have the highest frequency results either in <i>Cryptosporidium</i> samples or negative samples in order of retention times. Compound; 1-propanol was absent in <i>Cryptosporidium</i> samples but present in 11 control samples. Compounds; cyclopentane was present in 7 <i>Cryptosporidium</i> samples and only one in control samples and 3-hydroxy-2-butanone was present in 10 <i>Cryptosporidium</i> samples and 5 in control samples, these compounds have p-values of less than 0.05.	161
Table 1A. Demographic data of symptomatic children from Karonga in relation to <i>Giardia</i> and <i>Cryptosporidium</i> positive by PCR and DFA	185
Table 1B. Demographic data of symptomatic children from Lilongwe in relation to <i>Giardia</i> and <i>Cryptosporidium</i> positive by PCR and DFA	186
Table 1C. Demographic data of symptomatic children from Mangochi in relation to <i>Giardia</i> and <i>Cryptosporidium</i> positive by PCR and DFA	187
Table 2A. The distribution of <i>Giardia</i> assemblage A and B from Malawi based on the different loci and mixed infection	188
Table 2B. The distribution of <i>Giardia</i> assemblage A and B from Cambodia showing all positive results by PCR and based on the different loci and mixed infection	193
Table 2C. Demographic data of symptomatic children from Cambodia in relation to the distribution of <i>Giardia</i> sub-assemblage B based on the different loci and mixed infection	200
Table 5A. Shows total number that each compound occurred in 11 <i>Cryptosporidium</i> samples and 11 control samples. It also shows their Chi-Square test results and p-values for each compound between <i>Cryptosporidium</i> and each control. Highlighted compounds show compounds which are biologically significant in the identification of Cryptosporidiosis.	229

LIST OF ABBREVIATIONS

AFS	Acid Fast Stain
AHC	Angkor Hospital for Children
AIDS	Acquired Immune Deficiency Syndrome
AMDIS	Automated Mass Spectral Deconvolution and Identification System
ART	Antiretroviral Treatment
ASH	Allele Sequence Heterozygosity
<i>Bg</i>	<i>B-giardin</i> gene
Bgf/Bgr	Beta-giardin forward Primer/ Beta-giardin reverse Primer
BLAST	Basic Local Alignment Search Tool
BP	Base Pair
CAR/PDMS	Carboxen/Polydimethylsiloxane
CLUSTAL X	Program allowing multiple alignment of nucleic acid and protein sequences
COWP	<i>Cryptosporidium</i> Oocyst Wall Protein
°C	Degree Celsius
DEC	Diarrhoeagenic <i>Escherichia Coli</i>
DFA	Direct Fluorescence Assay
DNA	Deoxyribonucleic Acid
DNADIST	Program to compute distance matrix from nucleotide sequences
DNASP v3	DNA Sequence Polymorphism
dNTPs	Deoxy Nucleotide Triphosphates
EDTA	Ethylenediaminetetraacetic acid
EFL-alpha	Elongation Factor 1 – alpha
EI	Electron ionization
EIA	Enzyme Immune Assay
ELISA	Enzyme linked Immunosorbent Assay
GC	Gas Chromatography
GCMS	Gas Chromatography Mass Spectrometry
<i>Gdh</i>	Glutamate Dehydrogenase
GP60	Glycoprotein 60
HIV	Human Immunodeficiency Virus
HSP70	70kda Heat Shock Protein
IUPAC	International Union of Pure and Applied Chemistry

LLE	Liquid-liquid Extraction
MEGA	Molecular Evolutionary Genetics Analysis
MgCl ₂	Magnesium chloride
ML	Maximum Likelihood
μl	Microlitre
MLG	Multi-locus Genotype
MLST	Multi-locus Sequence Typing
MLT	Multi-locus Typing
mM	Millimolar
μM	Micromolar
MN	Mini-Satellite
MP	Maximum Parsimony
MS	Micro-Satellite
MZN	Modified Ziehl Neelsen
NGO	Non Governmental Organisation
NHS	National Health Service (UK)
NIST	The National Institute of Standards and Technology
NJ	Neighbour Joining
NMC	The national Centre for Parasitology, Entomology and Malaria Control
<i>P</i>	p-value
%	Percent
PCR	Polymerase Chain Reaction
PHYLP	Phylogeny Inference Package
PLS-DA	Partial Least Squares – Discriminant Analysis
Pmol	Pico mole
QECH	Queen Elizabeth Central Hospital
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid
RT	Retention Time
SDH	Sotr Nikom District Hospital
SEA	South East Asia
SNPs	Single nucleotide polymorphisms
SPME	Solid phase microextraction

spp.	Several species
SPSS (Version 7)	Statistical Package for the Social Science
<i>SspI</i>	<i>Sphaerotilus</i> species
SSU	Small Sub-Unit
STH	Soil Transmitted Helminths
TLC	The Lake Clinic
<i>Tpi</i>	Triosephosphate Isomerase
UK	United Kingdom
U	Unit
VOC's	Volatile Organic Compounds
<i>VspI</i>	<i>Vibrio</i> species
WHO	World Health Organisation

CHAPTER ONE

GENERAL INTRODUCTION

1. The global importance of intestinal parasites

Intestinal parasitic infections (helminths and protozoa) are of major concern to public health bodies throughout the world (WHO, 2000). At any one time, about 25% of the world's population suffer from one or more parasitic infections, which are endemic throughout the world, particularly in tropical and subtropical countries. It is estimated that about 3.5 billion people are infected with intestinal parasites and nearly 450 million suffer from clinical morbidity (WHO, 2000). The protozoan parasites are responsible for more gastrointestinal disorders than the helminths, particularly in developing countries.

G. duodenalis is the most prevalent protozoan parasite worldwide infecting about 200 million people (Arora and Arora, 2005). *G. duodenalis* is known to infect 2% to 5% of populations in developed countries and 20-40% in developing countries, the majority of which are children (Ali and Hill, 2003, Wani *et al.*, 2010). Global statistics on prevalence of *Cryptosporidium parvum* shows that it infects 2-50% of population worldwide (WHO, 2000). In Asia and Africa, the infection rate ranges from 5-10%.

Giardia duodenalis and *C. parvum* are major causes of diarrhoeal disease in humans worldwide. Both *Giardia* and *Cryptosporidium* have life cycles which are suited to waterborne and foodborne transmission. *G. duodenalis* is cyst-forming whilst *C. parvum* is intracellular and forms spores which are passed out with faeces (Smith *et al.*, 2006; Xiao *et al.*, 2004). These organisms are also found in various animals and zoonotic transmission is thought to occur (Feng & Xiao, 2011; Huang & White, 2006). In developing countries, detailed study of the relative importance of zoonotic transmission of *Giardia* and *Cryptosporidium* infections has not been undertaken (Mak, 2004). *G. duodenalis* cysts and *Cryptosporidium* oocysts have been detected in various ground water sources whereas maintenance of the infection and their role in various community outbreaks has not been completely characterised (Graczyk & Fried, 2007; Mak, 2004). Diarrhoea, especially chronic diarrhoea is associated with patients suffering from AIDS. Gassama *et al.* (2001) studied

the enteropathogens of patients with HIV. Of those patients that were infected with a diarrhoea causing parasite, *Cryptosporidium* was found in 8.2% of patients and *G. duodenalis* was found in 4.9% of patients, this was from a total sample size of 121 patients.

In general, the prevalence of *Cryptosporidium* infection rates are higher in immunocompromised patients compared with immunocompetent hosts. In addition, there are higher prevalence rates in rural communities compared with urban areas (Mak, 2004).

Most research has been undertaken in developed countries using different methodologies to try and understand the routes of infection, whereas research in developing countries has been limited (Snelling *et al.*, 2007).

The documented prevalence of a parasite depends upon the method of detection. Most studies on the prevalence of Giardiasis and Cryptosporidiosis in developing countries were based on conventional microscopy (Mak, 2004; Smith *et al.*, 2006). These results in developing countries might be limited due to a lack of laboratory resources, such as modern equipment and the quality materials necessary to carry out the research. This can affect the accuracy of the estimates of infection rates of both parasites in developing countries (Koch *et al.*, 1985; Mak, 2004; Snelling *et al.*, 2007).

1.2 Introduction to *Giardia duodenalis* and Giardiasis

Giardia duodenalis (syn. *Giardia intestinalis*, *Giardia lamblia*) is a flagellated protozoan parasite that causes giardiasis in both humans and animals. *Giardia* is also a common protozoan parasite of domestic animals, pets, livestock, birds, and domestic sheep and cows (Thompson, 2004; Thompson & Monis, 2004a; Xiao & Fayer, 2008), in addition to wildlife such as alpaca, banteng, bison, camel, deer, gayal, goat, horse, llama, yak, reindeer, water buffalo and mule (Appelbee *et al.*, 2005; Feng & Xiao, 2011).

1.2.1 Biology and life cycle of *Giardia*

Giardia is comprised of two stages: trophozoite and cyst. The infective stage of the parasite. The cyst are normally about 12-18 µm in length, 6-8 µm in width and 2-4 µm thick. They are encysted when released into the feces and is immediately infectious (Huang & White, 2006). The

trophozoites are binucleate microorganisms with a pear shape and convex dorsal surface. The trophozoites consist of 4 flagella pairs arranged in bilateral symmetry, with an adhesive disc located in the anterior half of the ventral surface. There are two median bodies transversely passed in the mid body (claw-hammer shaped) and flagella axonemes (axostyle) longitudinally located between nuclei. Trophozoite multiplication and encystation are stimulated by bile salts and intestinal mucus. The trophozoite stage is responsible for producing clinical disease in humans (O'Handley *et al.*, 2001). The cysts have an oval shape with a smooth wall (colourless and refractile). They are 8-14 μm in length and 6-10 μm in width. The mature cyst contains four nuclei. There are two median bodies transversely passed in the mid body (crescent-shaped), the axoneme located pass the long axis of the cyst. They are environmentally resistant and at this stage are responsible for disease transmission to another host. Cysts may remain infectious for several months in cool, damp areas and rapidly accumulate in, moist environmental conditions. They have also been detected in natural surface waters. They are also able to survive standard concentrations of chlorine used in water purification systems.

Infection occurs after cysts are ingested. When ingested by the host, cysts excyst in the duodenum, releasing the trophozoites. This marks the beginning of the life cycle. After ingestion, mature cysts in the small intestine release trophozoites through a process called excystation. The trophozoites release from each cyst and multiply asexually by longitudinal binary fission. The distribution of trophozoites within the intestine varies with host and diet (Greene *et al.*, 1990). The multiplication continues and the cysts are passed with stools, and can then be transmitted to new hosts. In cases where patients have severe diarrhea, some trophozoites may be detected. Cysts are able to survive exposure to gastric acid; gastric acid may actually booster excystation (Barr, 1998).

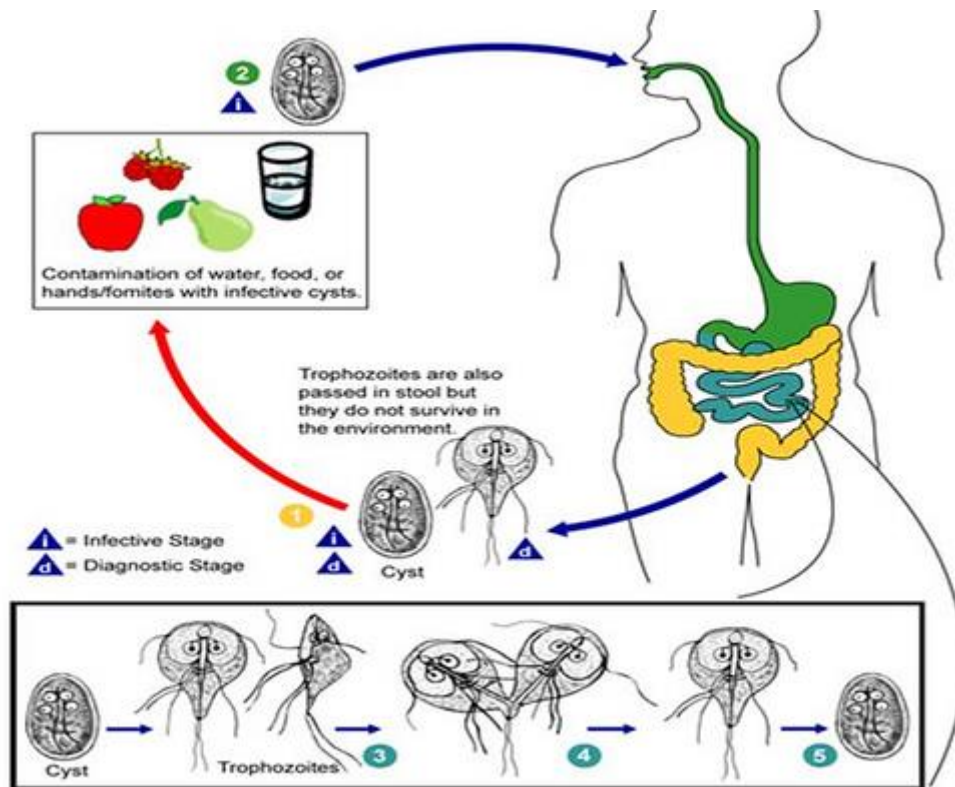


Figure 1.1 The life cycle of *G. duodenalis*. The cysts when passed from faeces, can survive in the environment for many months (1). The cysts from contaminated water or food are ingested by the host and then excysted in the duodenum (2). Trophozoites are released from each cyst (3). The trophozoites multiply asexually by longitudinal binary fission (4). The multiplication continues and the cysts are passed with stools, and can then be transmitted to new hosts (5). Picture taken from <http://www.dpd.cdc.gov/dpdx/cryptosporidiosis>.

1.2.2 Transmission of *Giardia*

Transmission of *Giardia* can be direct, from host to host or can be spread indirectly by water, by faecal-oral contamination with cysts or accidentally through the ingestion of contaminated food or water. Giardiasis has been shown to be sexually transmitted between male homosexuals. Oral, genital and digital contact with the anus can aid in the transmission of *G. duodenalis*. (Tessier and Davies, 1999).

Giardia is the most common cause of epidemic waterborne diarrheal disease (Homan & Mank, 2001). *Giardia* can be a cause of recreational water outbreaks of gastroenteritis in swimming

pools, water parks, hot tubs and spas. Infective cysts can also be found in contaminated soil and other objects of infection (Berger and Marr, 2006; Olson *et al.*, 2004)

The most effective method to kill chlorine-resistant *Giardia* cysts is boiling water and making water filtration the key to effective water treatment (Harter *et al.*, 1984). Humans who have voluntarily agreed to be infected have shown that an infectious dose can be as little as 10 cysts (Leber, and Novak-Weekley, 2007; Rendtorff, 1954). Protozoan contamination of water or food within communities is the cause of many outbreaks (Mak, 2004).

1.2.3 Taxonomy and genotyping of *Giardia*

There is only one species, *G. duodenalis* that is recovered from humans and most other mammals (Adam, 2001). Isolates of *G. duodenalis* are classified into seven common assemblages A, B, C, D, E, F, and G (Meloni *et al.*, 1995; Monis *et al.*, 2003). These classifications are based on the characterisation of the triosephosphate isomerase (*tpi*) (Amar *et al.*, 2002; Bertrand *et al.*, 2005), surface protein genes, glutamate dehydrogenase (*gdh*) (Bertrand *et al.*, 2005; Read *et al.*, 2004), β -*giardin* (*bg*) (Lalle *et al.*, 2005; Trout *et al.*, 2003), small subunit (SSU) rRNA and other genes (Hopkins *et al.*, 1997; Read *et al.*, 2004; Sulaiman *et al.*, 2003), elongation factor 1-alpha (EFL-alpha) (Traub *et al.*, 2004) and other catabolic enzyme genes, isoenzyme electrophoresis (allozyme analysis), and sequencing analysis (Homan & Mank, 2001; Thompson *et al.*, 2000).

Assemblages A infect humans and a wide range of other hosts such as cattle, dogs, cats, and also wild mammals (slow loris, livestock, deer, beavers, muskrats, voles, guinea pigs and ferrets). Moreover, the assemblage A is divided into two major sub-assemblages (I and II) and the less common III. Assemblage B is isolated from human and other mammals such as siamang, slow loris, livestock, chinchillas, dogs, beavers, muskrats, voles, rats and marmosets. However, there are many subtypes found in the assemblage B (Geurden *et al.*, 2009a; Geurden *et al.*, 2009b). Assemblages C, D, E, F, and G appear to be restricted to cattle, companion animals, and rodents (Sulaiman *et al.*, 2003). Assemblage C and D is found in dogs and coyotes. Assemblage E is isolated from alpaca, cattle, goats, pigs and sheep. Assemblage F and G are considered to be specific to cats and rats, respectively (Thompson, 2004). Based on trophozoite morphology, there are 6 *Giardia*

species as follows *G. duodenalis* infects mammals, *G. agilis* infects amphibians, *G. muris* infects rodents, *G. psittaci* and *G. ardeae* infect birds, *G. microti* infects the prairie vole. Research suggests that diverse species and genotypes show different levels of infectivity and virulence in both humans and animals in experimental studies (Okhuysen *et al.*, 1999).

1.2.4 Zoonotic transmission

Cattle are susceptible to transient infection with zoonotic genotypes, in which the frequency of transmission with the livestock genotype is high and competition is likely to occur. Domestic dogs are susceptible to infection with host-adapted *G. duodenalis* assemblage C or *G. canis* and by zoonotic transmission. Zoonotic transmission of *Giardia* genotypes has been shown to occur where humans and dogs live closely together (Traub *et al.*, 2004).

In particular, *Giardia* isolates that are morphologically identical to *G. duodenalis*, the occurrence of isolates in wildlife has been the single most important factor implicating *Giardia* as a zoonotic agent. There is little evidence to support the role of wildlife (animals that are neither kept domestically or farmed) as a source of disease in humans. Scant evidence exists to corroborate the idea that in waterborne outbreaks these infections are likely to be the original source of contamination, particularly that the source could be traced to aquatic mammals. (Thompson, 2004; Thompson & Monis, 2004a; Thompson & Monis, 2004b). Animals are likely to have become infected from water contaminated with human faecal material. The source of *Giardia* infection in beavers was likely to be of human origin, the few studies that have genotyped *Giardia* of beaver origin and confirmed previous suggestions (Appelbee *et al.*, 2005; Sulaiman *et al.*, 2003).

1.2.5 Symptoms and Clinical features of Giardiasis

The major symptom of acute giardiasis is diarrhoea. The incubation period of infection is about 9-15 days. The acute stage often begins with an uncomfortable feeling in the stomach followed by nausea, vomiting and anorexia (Wolfe, 1992). Otherwise, fever and chills may occur. These symptoms are followed by watery, explosive foul-smelling diarrhoea, abdominal pain, passage of foul gas and belching (Hanevik *et al.*, 2007). This stage is generally 3-4 days. However, symptoms may persist for months if left untreated. Malabsorption due to chronic *Giardia* infection has also

been reported. After several days of Giardiasis, weight loss and dehydration can often occur (Adam, 2001; Greene *et al.*, 1990). The diarrhoea can be mild and produce semi-solid stools, or it can be intense with greasy stools (Wolfe, 1992). In general, children become less ill than adults and frequently develop asymptomatic infection. It has been demonstrated that as many as 50% of infections are asymptomatic (Flanagan, 1992). Clinical presentation ranges from asymptomatic cyst passage to chronic diarrhea, malabsorption, severe weight loss, and malnutrition (Adrubbo and Peura, 2002).

Giardiasis is considered a zoonotic disease and this protozoan causes a frequently self-limited clinical illness identified by diarrhea, bloating, abdominal cramps, weight loss, and malabsorption (Hanevik *et al.*, 2008). However, asymptomatic giardiasis occurs commonly in developing countries. Giardiasis is correlated with a spectrum of symptomatology, including acute and chronic malabsorptive, allergic manifestations and childhood failure to thrive.

Giardiasis may present with a wide variety of clinical manifestations. The majority (60–80%) of infected individuals have few or no symptoms (Flanagan, 1992; Farthing, 1992; Walterspiel and Pickering, 1994). Asymptomatic giardiasis may be epidemiologically more significant, as asymptomatic individuals are less likely to be detected or to seek treatment and therefore, are more likely to serve as carriers of the disease (Meyer, 1980). Symptomatic giardiasis is characterized by the acute onset of diarrhea, abdominal pain or cramps, and flatulence. Diarrhea is the predominant symptom, occurring in 90% of symptomatic individuals (Gorski, 1985). Patients pass approximately five loose or semifformed stools per day for 2–3 days. The episodes of diarrhea may be interposed with periods of normal bowel movements or constipation. The flatulence is characterized by a sulfur odor, which is frequently emitted from the breath, stool, or both. Untreated, these symptoms may persist for months and lead to overt malabsorption and weight loss. In infants and children, prolonged infection may lead to impaired growth and development (Farthing, 1992). Fever, chills, severe abdominal pain, and bloody stool are uncommon symptoms of *Giardia* infection, and, if present, other diagnoses must be considered (Gorski, 1985; Babb, 1995). Extraintestinal manifestations of giardiasis are rare but include maculopapular rash, urticaria, dermatitis, arthritis, and biliary tract disease. Acute giardiasis must be differentiated from other acute diarrheal illnesses, such as viral or bacterial gastroenteritis and food poisoning. Similarly, the differential

diagnosis of chronic giardiasis includes inflammatory bowel disease, irritable bowel syndrome, celiac disease, and other infections caused by organisms such as *Isospora* and *Strongyloides*.

1.2.6 Pathogenesis of Giardiasis

Giardia duodenalis infection causes chronic diarrhoea and malabsorption however this process of symptomatic infection is still not fully understood. Some degree of mucosal damage is common which associates with animal models of infection (Buret *et al.*, 1991; Buret *et al.*, 1992; Dagci *et al.*, 2002). Other mechanisms are thought to include direct physical damage, release of parasitic products such as proteinases or lectin and mucosal inflammation associated with T cell activation and cytokine release (Katelaris and Farthing, 1992; Faubert, 2000). Other methods of malabsorption may include associated bacterial overgrowth and bile salt deconjugation, bile salt uptake by the parasite with depletion of intraluminal bile salts, and inhibition of pancreatic hydrolytic enzymes (Katelaris and Farthing, 1992; Buret *et al.*, 1991). Therefore, it can be concluded that not one single mechanism explains the diarrhoea and malabsorption caused by *Giardia*, and that it can be viewed as a process of many factors (Nain *et al.*, 1991; Katelaris and Farthing, 1992).

1.2.7 Treatment of giardiasis

Drugs commonly used for the treatment of giardiasis are of the 5-nitroimidazole group such as Metronidazole, Tinidazole, Ornidazole and Secnidazole (Robertson *et al.*, 2010). However, Metronidazole is commonly used as the drug of choice (John & Petri, 2006). Other groups of drugs that can be used are as follows, Acridine analogs (Quinacrine), Nitrofurans (Furazolidone), Aminoglycosides (Paromomycin), Benzimidazoles (albendazole, mebendazole), 5-Nitrothiazoles (nitazoxanide) (Robertson *et al.*, 2010).

Mode of action of each drug

Metronidazole – a synthetic antibacterial and antiprotozoal agent of the nitroimidazole class.

Mode of action – anaerobic bacteria intracellularly reduce Metronidazole to its active form. The reduced Metronidazole is then able to bind to DNA, disrupting its structure and inhibiting nucleic acid synthesis resulting in cell death (Brayfield, 2014).

Tinidazole – the mechanism by which Tinidazole is effective against *Giardia* is not fully known. However chemically reduced Tinidazole has been demonstrated to release nitrite and cause damage to purified bacterial DNA in laboratory conditions (Edwards, 1993).

Ornidazole – taken orally, Ornidazole is absorbed into the bloodstream and taken up by tissues. It interacts with the Protozoan DNA and blocks the formulation of nucleic acids, thus killing the protozoa (Rutgeerts *et al.*, 2005).

Secnidazole – Secnidazole is diffused into the parasite and then intracellularly reduced by low oxidation-reduction potential ferredoxin this then results in damage to the protozoa DNA (Gillis & Wiseman, 1996).

Furazolidone - believed that the free radicals it produces bind to DNA and induce crosslinks affecting the ability to make proteins (Machado *et al.*, 2008).

Aminoglycoside antibiotics (including Paramomycin) – effective by binding to the Ribosomes. Thus, inhibiting the translocation of RNA and negating the ability to synthesise proteins (Davidson *et al.*, 2009).

Benzimidazole – interferes with the normal interaction of ubiquinone and the cyt bc1 complex which then disrupts the mitochondrial electron transport (Huynh *et al.*, 2005).

Albendazole and Mebendazole – degenerate the tegument and intestinal cells of the worm by binding to the tubulins colchicine sensitive side. This inhibits the assembly of microtubules leading to impaired uptake of glucose in the larval and adult stages. Diminished energy leads to the parasites death (Gardner & Hill, 2001).

Nitazoxanide – works by interfering with the PFOR enzyme dependent electron transfer reaction. This reaction is essential for the metabolism of anaerobic energy by the parasite (Musher *et al.*, 2009).

1.2.8 Prevention of Giardiasis

It is not possible to prevent infection from *Giardia* by taking preventative medicine. Maintaining good standards of hygiene and following a number of precautions such as those listed below are

the best way to avoid getting infected or indeed passing the infection to others (CDC *Giardia*, 2011).

Hand washing – always wash your hands with soap (preferably anti-bacterial) and water particularly after the following activities; going to the toilet, changing nappies and before handling or preparing food. A number of hand sanitisers are available if soap and water isn't and these are easy to carry with you (CDC *Giardia*, 2011).

Drinking water – never drink untreated water from wild sources without ensuring it has been filtered or boiled for at least 10 minutes. Only drink and brush your teeth with bottled water in areas where water quality is likely to be compromised (Betancourt, 2004). Avoid ice in your drinks and avoid the consumption of raw fruits and vegetables (CDC *Giardia*, 2015). Water in swimming pools, even if chlorinated can contain *Giardia* parasites so avoid accidentally swallowing water in these areas and carefully monitor children (Exner and Gornik (2004).

Avoid spreading infection – if you have contacted Giardiasis, to help stop spreading the infection you must wash your hands regularly as described above, avoid handling or preparing food that could be eaten by family members, friends or colleagues. Do not share towels and stay away from work/school until you have been without symptoms for 48 hours (Workowski and Bolan, 2015).

Practice safe sex – always use a condom if you engage in anal or oral-anal sex. Wash hands after handling a used condom. Avoid kissing or licking as this can increase the risk of infection (Mayo clinic, 2015).

1.3 Introduction to *Cryptosporidium* and Cryptosporidiosis

There are more than 27 accepted *Cryptosporidium* species and a nearly 50 genotypes have been described in animals (Feng *et al.*, 2007a; Feng *et al.*, 2007b; Feng *et al.*, 2007c; Xiao *et al.*, 2004). There are at least seven species of *Cryptosporidium* are associated with human disease consists of *C. hominis*, *C. parvum*, *C. meleagridis*, *C. felis*, *C. canis*, *C. suis* and *C. muris*. Another species, *C. ubiquitum* can cause disease in humans (Chalmers *et al.*, 2009). Moreover, two genotypes of *Cryptosporidium* sp. (monkey and cervine) are also have been investigated in human faeces. *C. parvum* is a cause of diarrhoea in both immunocompromised and immunocompetent persons (Kjos *et al.*, 2005). Moreover, *C. parvum* is the major zoonotic *Cryptosporidium* species which causes neonatal diarrhoea in livestock, with consequent economic loss. Therefore, *C. parvum*

cryptosporidiosis in cattle is a major contributor to environmental contamination with oocysts. Wildlife can also harbour *C. parvum*, and molecular investigations into the species and genotypes isolated from wildlife indicate that they also harbour their own host-adapted species, which may not be infectious to humans. Development of molecular studies help determine the zoonotic potential of oocysts found as contaminants of water and food. Molecular advances have enabled a better understanding of the contributions of humans and cattle as reservoirs of infection. Using species-typing tools disease presentations, risk factors for infection, and differences in geographical and transient distribution have been characterised for *C. parvum* and *C. hominis*, the most commonly reported causes of human cryptosporidiosis. There is extensive genetic variation within the genus *Cryptosporidium*. Recently, phylogenetic studies found that *Cryptosporidium* species and genotypes form mostly two groups: those found primarily in the intestine and those in the stomach (Ryan and Xiao, 2004; Feng *et al.*, 2007a). Each group contains parasites of mammals, birds and reptiles. With few exceptions, most species and genotypes are host-adapted in nature, having a narrow spectrum of natural hosts. Thus, one *Cryptosporidium* species or genotype usually infects only a particular species or a group of related animals. The existence of host-adapted *Cryptosporidium* species or genotypes indicates that cross-transmission of *Cryptosporidium* among different groups of animals is probably limited. *C. parvum* has received the most attention in zoonotic transmission of cryptosporidiosis. This was largely due to the fact that *C. parvum* is a major human pathogen and was traditionally considered to infect all mammals. Genetic characterisations of *Cryptosporidium* specimens from various animals, however, have mostly failed to detect this parasite in wild mammals (Feng *et al.*, 2007a; Zhou *et al.*, 2004).

1.3.1 Biology and life cycle of *Cryptosporidium* spp.

The lifecycle begins with the ingestion of sporulated oocysts (Figure 1.2). The oocysts each contain four sporozoites which are found in the infected hosts faeces. Upon ingestion by a suitable host excystation (a) occurs, where upon the sporozoites are released and parasitize the epithelial cells (b, c) of the gastrointestinal tract. The ileum is infected and sporozoites penetrate individual epithelial cells in this region. The parasite asexually multiplies and oocysts (j, k) develop that sporulate in the infected host and are subsequently excreted. Because the oocysts are infective upon excretion immediate fecal/oral transmission can occur (Juranek, 2000; Figure 1.2).

The diagram illustrates the life cycle of *Toxoplasma gondii*, divided into environmental and internal stages.

Environmental Cycle:

- (1)** A person contaminates water with feces.
- (2)** The contaminated water is consumed by another person through drinking water or food, leading to infection.

Internal Cycle:

- (j)** A thick-walled sporulated oocyst exits the host.
- (a)** The oocyst is ingested.
- (b)** The sporozoite develops.
- (c)** The sporozoite develops into a trophozoite.
- (d)** The trophozoite can undergo an asexual cycle, becoming a Type I meront.
- (e)** The Type I meront develops into a merozoite.
- (f)** The merozoite can undergo a sexual cycle, becoming a Type II meront.
- (g)** The Type II meront develops into an undifferentiated gamont.
- (h)** The undifferentiated gamont develops into a microgamont.
- (i)** The microgamont develops into a macrogamont.
- (j)** The macrogamont and microgamont fuse to form a zygote.
- (k)** The zygote develops into a thin-walled sporulated oocyst, which can cause auto-infection.

1.3.2 Transmission of *Cryptosporidium* sp.

12

C. hominis in cattle in experimental (Giles *et al.*, 2001) and natural infections (Smith *et al.*, 2005) and other non-human hosts (Morgan *et al.*, 2000a), *C. hominis* is primarily an infection of humans. Nonetheless, *C. meleagridis*, *C. muris*, *C. suis*, *C. felis* and *C. canis* and the *Cryptosporidium* cervine and monkey genotypes also infect humans. Otherwise, person-to-person transmission is assumed for *C. meleagridis*, *C. muris*, *C. suis*, *C. felis* and *C. canis* and the *Cryptosporidium* cervine and monkey genotypes. Nevertheless, little is known about their zoonotic capability.

However, *Cryptosporidium* oocysts can pervade physical obstructions in water treatment processes, are disinfectant resistant and are extensively distributed in the surroundings (Graczyk & Fried, 2007). Many researchers demonstrated that *Cryptosporidium* oocysts have the highest potential for transmission through drinking water. In addition, *Cryptosporidium* has a low infectious dose for humans. As few as two to 10 oocysts can initiate an infection (Chen *et al.*, 2003). Oocysts are passed in stools and can contaminate water directly or indirectly. The removal of human and animal remains is a significant public health issue because prevention and control in many parts of the world has not been systematically evaluated. Transposition of *Cryptosporidium* by means of drinking and recreational water is well reported, as are outbreaks of cryptosporidiosis following consumption of unclean water (Karanis *et al.*, 2007; Savioli *et al.*, 2006; Smith *et al.*, 2003). Cryptosporidiosis is more common in summer time (between December and May) than in winter time (between June and November) in Peru (Bern *et al.*, 2002). Whereas the number of *Cryptosporidium* spp. infections was higher in the wet season than in the dry season in Zambia (Siwila *et al.*, 2011). Moreover, the *Cryptosporidium* infection associated with rainy season in Blantyre, Malawi (Peng *et al.*, 2003).

The role of companion animals in the transmission of human cryptosporidiosis is less important. It has been suggested for some time that dogs can be a significant source of human cryptosporidiosis (Enriquez *et al.*, 2001; Shukla *et al.*, 2006). Nevertheless, this was largely based on the observation of direct transmission of *C. parvum* from calves to humans and the wrong belief that *C. parvum* is responsible for cryptosporidiosis in all mammals.

1.3.3 Taxonomy and Genotyping of *Cryptosporidium* spp.

The main causative agents of human cryptosporidiosis are genotype of human (genotype 1) and bovine genotypes (genotype 2) of the species *C. parvum* (Fayer *et al.*, 2000a; Fayer *et al.*, 2000b). Researchers found at least seven *Cryptosporidium* species, (*C. hominis*, *C. canis* (*C. parvum* dog), *C. felis* (*C. parvum* cat), *C. meleagridis* (avian parasites), *C. muris* (murine parasites), *C. suis*, and *C. andersoni*) have been reported to cause human disease (Gatei *et al.*, 2002; Morgan *et al.*, 2000b; Morgan *et al.*, 2000c; Pedraza-Diaz *et al.*, 2001; Xiao *et al.*, 2001a). Moreover, it has been shown that at least six genotypes (cervine, skunk, chipmunk I, horse, rabbit genotype, and pig genotype II) can infect humans (Leoni *et al.*, 2006a; Leoni *et al.*, 2006b; Read *et al.*, 2004). However, *C. hominis* and *C. parvum* were responsible for most clinical cases. The difference between these two genotypes is that *C. hominis* is responsible for human to human transmission while *C. parvum* is responsible for animal to human transmission. Cattle have been described as major hosts for *C. parvum*, *C. bovis*, *Cryptosporidium* deer-like genotype and *C. andersoni*. Identification of the isolates at the subgenotype level will be useful for control of the cryptosporidial infections and for understanding the population structure of *C. parvum* and *C. hominis* genotypes. Only one or two *C. parvum* subtype families (IIa and IIc) were found in humans in India, Peru, New Orleans, Malawi, South Africa, Kuwait, and Portugal. In contrast, three to four subtype families from the five common *C. hominis* subtype families (Ia, Ib, Id, Ie, and If) were seen in humans in the same areas (Alves *et al.*, 2006; Gatei *et al.*, 2007; Sulaiman *et al.*, 2005). Zoonotic strains of *C. parvum* produced more severe infections in humans than the strains found only in humans (Okhuysen & Chappell, 2002).

1.3.4 Zoonotic transmission

Foods associated with oocyst contamination from non-human sources include raw vegetables, cold drinks made from contaminated water and shellfish which filter and retain viable oocysts. There are fewer foodborne outbreaks documented than waterborne outbreaks, this may be due to the lack of appropriate tools for detection and characterization of *Cryptosporidium* spp. Zoonotic transmission has been suggested for cryptosporidiosis outbreaks following the consumption of commercially obtained unpasteurised cow's milk and the consumption of non-alcoholic apple cider.

GP60 sub-genotyping has provided supportive evidence of association between oocyst contaminated apples, infected calves and contaminated water used to wash apples in previous outbreaks (Blackburn *et al.*, 2006). Neonatal calves are demonstrated to be an important source of zoonotic cryptosporidiosis in humans, though little is known about its transmission dynamics. Moreover, the potential reservoir hosts and transmission pathways for the novel species infecting humans are unclear (Fayer *et al.*, 2000a). Currently, DNA sequence analysis of the 60 kDa glycoprotein (GP60) gene has shown the complexity of *Cryptosporidium* transmission in endemic areas. Researchers have used highly discriminatory sub-typing techniques (GP60 gene), which are useful when tracking infection sources and examining the transmission dynamics of *C. parvum* (Abe *et al.*, 2006; Xiao *et al.*, 2006a; Xiao *et al.*, 2006b).

1.3.5 Symptoms and Clinical features of Cryptosporidiosis

The prepatent period, which is the interval between infection and the first appearance of oocysts in the feces, is usually 4 days (3 days in heavy infections) (Leder, 2016; Xiao *et al.*, 2004). Diarrhea, abdominal cramping and weight loss are clinical signs of this disease. Self limiting diarrheal disease lasts 7-14 days in immunocompetent persons (Petersen, 1995). In immunosuppressed hosts (AIDS/HIV, cancer patients and organ transplant patients) Cryptosporidiosis can be severe, debilitating and lead to death, with parasites sometimes found in the stomach, respiratory tract, biliary and pancreatic ducts (Gross *et al.*, 1986; Soave *et al.*, 1988).

Cryptosporidiosis can cause diarrhea in both immunocompromised individuals and immunocompetent hosts worldwide (Abdel-Hafeez *et al.*, 2012; Bajer *et al.*, 2008; Snelling *et al.*, 2007). Generally, young children and immunosuppressed patients are infected with this parasite. As well as those symptoms described above the infection can also cause loss of appetite, nausea, vomiting and fever. The symptoms of cryptosporidiosis usually appear within two to ten days after ingestion of the parasite.

Cryptosporidium spp. can cause asymptomatic infection (showing no symptoms) and diarrhea ranging from mild cases to severe enteritis (Jokipii and Jokipii, 1986; Checkley *et al.*, 1998). Asymptomatic infection often occurs in patients who are immunocompetent and immunodeficient. The incubation period is normally 7 to 10 days with a range of between 3 to 28 days. Oocyst

ingestion seems to be related to the infections duration (DuPont *et al.*, 1995). Patients can often suffer from fatigue, nausea and anorexia as well as abdominal pain and fever. Severe weight loss can result in patients with chronic diarrhea. Healthy people with no immunological conditions will normally recover in 10 to 14 days (Jokipii and Jokipii, 1986). Blood in the faeces is rare. Infants can suffer from poor growth patterns as a result of lack of nutrition when diarrhea is persistent. Immunocompromised patients the infection can be more severe and lead to muscle and tissue wasting. Different subtypes of *Cryptosporidium* spp. such as *C. hominis* and *C. parvum* can show different manifestations such as vomiting in addition to the diarrhea (Cama *et al.*, 2007). Other observed manifestations in AIDS patients have included cholecystitis, cholangitis, hepatitis, pancreatitis and occasionally the respiratory tract (Gross *et al.*, 1986; Soave *et al.*, 1988). Furthermore, children frequently have multiple episodes of cryptosporidiosis (Cama *et al.*, 2006). However, the likelihood of clinical illness decreases with increased infection episodes in developing countries (Cama *et al.*, 2008).

1.3.6 Pathogenesis of Cryptosporidiosis

The parasite is located in the brush border of the epithelial cells in the jejunum section of the small intestine. Attached sporozoites are enveloped by the epithelial cell membrane, making them intracellular but extracytoplasmic (Ryan *et al.*, 2004). The attached parasite can cause damage to the microvilli (Brooks *et al.*, 2004). Eventually this can lead to loss of the surface epithelium and the microvillous border being displaced, resulting in villous atrophy, blunting, crypt cell hyperplasia and mononuclear cell infiltration in the lamina propria (Chen *et al.*, 2003). What causes diarrhea to occur is not fully understood. It is known that the organism causes diarrhea that is associated with malabsorption. The fact that infection is intracellular means it is likely that it interferes with intestinal absorption and secretion. Given the large volume and watery nature of the diarrhea resembling secretory, this may indicate an enterotoxin as a specific mechanism, although no such toxin has yet been isolated (Ryan *et al.*, 2004).

1.3.7 Treatment of *Cryptosporidium*

Treatment is not generally prescribed for healthy patients and recovery generally occurs within two weeks. However, consideration should be given to maintaining a good intake of fluids and electrolytes to maintain good bodily function. Nitazoxanide can be prescribed to those with

healthy immune systems, but there are doubts about the effectiveness of this drug in those patients who are immunocompromised (Cabada *et al.*, 2015).

Nitazoxanide is an antiparasitic effective for the treatment of *Cryptosporidium* and *Giardia*. It works by interfering with the PFOR enzyme dependant electron transfer reaction. This reaction is essential for the metabolism of anaerobic energy by the parasite Abboud *et al.*, 2001; Gascon *et al.*, 1990).

For those patients whose immune systems are compromised (including those with HIV/AIDS) a range of treatments are available;

Loperamide – slows intestinal mobility by inhibiting the actions of the muscle walls of the intestines (Katzung, 2004; Lemke *et al.*, 2008).

Paramomycin - effective by binding to the Ribosomes. Thus, inhibiting the translocation of RNA and negating the ability to synthesise proteins (Morch *et al.*, 2008)

Azithromycin – similar mode of action to Paramomycin by binding to the bacterial ribosomes it inhibits protein synthesis in the bacteria (Greenwood, David (2008). The American Society of Health-System Pharmacists, 2015).

Rifaximin – inhibits RNA synthesis by binding to the bacterial RNA polymerase thus preventing transcription (PharmD & Benjamin Barner, 2010; DrugBank, 2017)

Rifabutin – particularly effective in immune suppressed HIV patients, prevents RNA synthesis thus causing bacterial cell death (Hamilton, Richart, 2015).

Spiramycin – inhibits translocation by binding to bacterial ribosomal subunits (WHO, 2015).

Clarithromycin – binds to ribosomal RNA through the bacterial cell wall blocking translocation and protein synthesis (The American Society of Health-System Pharmacists, 2015; Greenwood, 2008).

Octreotide – binds to somatostatin receptors eventually leading to inhibition of adenylyl cyclase (Allen *et al.*, 2000).

Atovaquone – is thought to inhibit electron transport, resulting in an inhibition of enzymes ultimately preventing nucleic acid and ATP synthesis (Farnert *et al.*, 2003; Krause *et al.*, 2000).

1.3.8 Prevention of Cryptosporidiosis

As with Giardiasis, maintaining good standards of hygiene and following a number of precautions such as those listed below are the best way to avoid getting infected or indeed passing the infection to others (Holmberg *et al.*, 1998; Fichtenbaum *et al.*, 2000).

Hand washing – always wash your hands with soap (preferably anti-bacterial) and water particularly after the following activities; going to the toilet, gardening where you have been in contact with the soil, changing nappies and before handling or preparing food. A number of hand sanitisers are available if soap and water isn't and these are easy to carry with you. You should also wash your hands after contact with animals (domestic, farm or wild). Drinking water – never drink untreated water from wild sources without ensuring it has been boiled for at least 10 minutes. Filtering water alone is not always effective as *Cryptosporidium* can resist chemical disinfection procedures. Only drink and brush your teeth with bottled water in areas where water quality is likely to be compromised. Avoid ice in your drinks and avoid the consumption of raw fruits and vegetables. Water in swimming pools, even if chlorinated can contain *Cryptosporidium* parasites as they are resistant to chlorination, so avoid accidentally swallowing water in these areas and carefully monitor children (Korich *et al.*, 1990). Avoid unpasteurized milk – avoid this and other dairy products made from unpasteurized milk such as certain cheeses (Chen *et al.*, 2003). Be careful around animals – wash hands after contact with animals or if you have been in areas where the presence of animal faeces is likely, wear disposable gloves and wash hands if there is a need for the handling of animal faeces, such as picking up dog fouling (Brooks *et al.*, 2004). Avoid spreading infection – if you have contacted *Cryptosporidium*, to help stop spreading the infection you must wash your hands regularly as described above, avoid handling or preparing food that could be eaten by family members, friends or colleagues. Do not share towels and stay away from work/school until you have been without symptoms for 48 hours (Murray *et al.*, 2005). Practice safe sex – always use a condom if you engage in anal or oral-anal sex. Wash hands after handling a used condom. Avoid kissing or licking as this can increase the risk of infection (Holmberg *et al.*, 1998; Fichtenbaum *et al.*, 2000). HIV infected patients – should take time to fully understand the ways to avoid being infected with *Cryptosporidium* and be aware of the risks associated with any of the above activities (Holmberg *et al.*, 1998; Fichtenbaum *et al.*, 2000).

1.3.9 *Cryptosporidium* species/genotypes in humans

There is evidence that more than one *Cryptosporidium* species is involved in human disease (Hunter & Thompson, 2005). *C. parvum* and *C. hominis* are the most common species causing human disease (Hunter & Thompson, 2005; Slapeta, 2006). Although, *C. hominis* and *C. parvum* are most commonly responsible for the majority of human infections. In some areas, *C. meleagridis* infection rates are as high as *C. parvum* (Cama *et al.*, 2008; Cama *et al.*, 2007). In addition, other *Cryptosporidium* genotypes are occasionally found in humans such as *Cryptosporidium* cervine, horse, rabbit, skunk and chipmunk I genotypes (Xiao & Fayer, 2008; Xiao & Feng, 2008; Xiao and Ryan, 2008). Nonetheless, these are rare in humans. The distribution of *C. parvum* and *C. hominis* in humans differs in geographic regions. In the Middle East, *C. parvum* is the dominant species in humans (Meamar *et al.*, 2007; Pirestani *et al.*, 2008; Tamer *et al.*, 2008).

In developing countries, *C. hominis* is usually the predominant species in humans (Araujo *et al.*, 2008; Gatei *et al.*, 2008; Jex & Gasser, 2008). While *C. parvum* and *C. hominis* are common in humans in European countries (Bajer *et al.*, 2008; Chalmers *et al.*, 2009; Savin *et al.*, 2008; Wielinga *et al.*, 2008; Zintl *et al.*, 2009). However, geographic variations can also occur within a country for the distribution of *C. parvum* and *C. hominis*. For example, in Peru, researchers found that there were no significant differences in the distribution of *Cryptosporidium* species or genotypes between children and HIV+ persons, indicating that there is no infection with zoonotic species/genotype in immunocompromised persons (Cama *et al.*, 2008; Cama *et al.*, 2007). *C. parvum* transmission in humans in developing countries appears largely anthroponotic. The zoonotic IIa subtypes are rarely seen in humans in these regions. The anthroponotic IIc subtype family is responsible for most human *C. parvum* infections in these areas (Table 1.1 and Table 1.2). In some regions such as Lima, Peru and Kingston, Jamaica, the IIc subtype family is the only *C. parvum* in humans, while in other developing countries such as India, Malawi, Uganda and Kenya, some unusual *C. parvum* subtype families such as IIb and IIe, are also seen in humans, which have never been seen in animals (Table 1.1 and 1.2). The anthroponotic nature of IIc still needs the support of results from animal studies. Only a few *C. parvum* isolates from calves in these areas have been done on GP60 subtyping (Feng *et al.*, 2007c). Nevertheless, the predominance of *C. hominis* in developing countries is also in agreement with the anthroponotic transmission of *C. parvum* IIc in

these areas. This is because both are transmitted through anthroponotic routes. A study on MLST was conducted in Jamaica to confirm the existence of human-adapted *C. parvum* (Gatei *et al.*, 2008).

Table 1.1 Occurrence of *Cryptosporidium* spp. in immunocompetent individuals in developing countries

Country	Sample size	Age (Maturity)	<i>C. parvum</i>	<i>C. hominis</i>	<i>C. meleagridis</i>	<i>C. felis</i>	<i>C. canis</i>	PCR types (gene targets)	References
Nepal	2	Not known	-	2	-	-	-	GP60	Wu <i>et al.</i> (2003)
Kenya	9	All	-	9	-	-	-	18s rDNA	Gatei <i>et al.</i> (2003)
Malawi	9	Children	3	6	-	-	-	18s rDNA	Gatei <i>et al.</i> (2003)
Malawi	43	Children	2	41	-	-	-	18s rDNA, hsp70, GP60	Peng <i>et al.</i> (2003)
Peru	118	Children	20	76	10	4	9	18s rDNA, hsp70	Xiao <i>et al.</i> (2001a)
Brazil	7	Children	-	7	-	-	-	18s rDNA	Gatei <i>et al.</i> (2003)

Table 1.2 Occurrence of *Cryptosporidium* spp. in immunocompromised individuals in developing countries

Country	Sample size	<i>C. parvum</i>	<i>C. hominis</i>	<i>C. meleagridis</i>	<i>C. felis</i>	<i>C. canis</i>	<i>C. muris</i>	<i>C. suis</i>	PCR types (gene targets)	References
Vietnam	3	-	3	-	-	-	-	-	18s rDNA	Gatei <i>et al.</i> (2003)
Thailand	29	-	24	3	1	-	1	-	18s rDNA	Tiangtip and Jongwutiwes <i>et al.</i> (2002)
Thailand	34	5	17	7	3	2	-	-	18s rDNA	Gatei <i>et al.</i> (2002)
Peru	299	34	204	38	10	12	-	1	18s rDNA	Cama <i>et al.</i> (2003)
Kenya	6	1	4	1	-	-	-	-	18s rDNA, hsp70, acetyl-CoA	Morgan <i>et al.</i> (2000c)
Kenya	24	8	14	1	-	-	1	-	18s rDNA	Gatei <i>et al.</i> (2003)
Malawi	2	1	1	-	-	-	-	-	18s rDNA	Gatei <i>et al.</i> (2003)

1.3.10 *Cryptosporidium hominis* infections

The proportion of human infections attributable to *C. hominis* differs between developed and developing countries. This may be due to differences in the disease burden attributable to zoonotic transmission. Moreover, developed and developing countries have some differences in molecular epidemiology of *C. hominis*. In developing countries, *C. hominis* infections in humans is reflected by the occurrence of multiple subtype families and multiple subtypes within families Ia and Id. For

example, researchers found 3–4 *C. hominis* subtype families in humans in India, Peru, Kenya, Malawi and South Africa, and there were many subtypes within *C. hominis* subtype families Ia and Id in one endemic area (Ajjampur *et al.*, 2007; Cama *et al.*, 2008; Cama *et al.*, 2007; Gatei *et al.*, 2007; Muthusamy *et al.*, 2006). *C. hominis* heterogeneity in developing countries is high and likely an indicator of intensive and stable cryptosporidiosis transmission in these areas. Otherwise, four *C. hominis* subtype families, Ia, Ib, Id and Ie, are seen in humans in these regions as well. Nevertheless, there are geographic differences in the distribution of them. For example, all the four common subtype families were seen in HIV+ adults and children in India, Peru, Madagascar and Malawi (Ajjampur *et al.*, 2007; Areeshi *et al.*, 2008; Cama *et al.*, 2008; Cama *et al.*, 2007; Gatei *et al.*, 2007; Muthusamy *et al.*, 2006). Conversely, in some South African children, scientists found subtype family If instead of Ie (Leav *et al.*, 2002). In Jamaica, the major *C. hominis* subtype family that was seen in HIV+ persons was Ib subtype family, with only a few other cases caused by other subtype families (Gatei *et al.*, 2008). Within each subtype family, only one subtype is frequently seen in certain areas but not in others. For example, there are only two common subtypes within the *C. hominis* subtype family Ib: IbA9G3 and IbA10G2. The former is generally seen in India, Kenya and Malawi, whereas the latter is commonly seen in Jamaica, Peru, South Africa and Botswana (Cama *et al.*, 2008; Cama *et al.*, 2007; Gatei *et al.*, 2008; Gatei *et al.*, 2007; Peng *et al.*, 2003). Several rare subtypes, IbA19G2, IbA20G2, IbA21G2 and IbA23G2, were the most common *C. hominis* subtypes in Shanghai and Tianjing, China (Feng *et al.*, 2009; Peng *et al.*, 2001). One unusual Ib subtype, IbA13G3, was seen in Peru (Cama *et al.*, 2007). In many developing countries, humans with the subtype family Ie are mostly infected with IeA11G3T3, with the exception of Kingston, Jamaica and Shanghai, China, where IeA12G3T3 is seen (Feng *et al.*, 2009; Gatei *et al.*, 2008). By using MST and MLST tools, specimens from different countries were genetically compared and found that this geographic segregation of *C. hominis* subtypes became much more obvious (Gatei *et al.*, 2008; Gatei *et al.*, 2006; Tanriverdi *et al.*, 2008). The clinical significance of various *C. hominis* subtypes in humans is still not clear. In Lima, Peru, infections with *C. hominis* Id were significantly associated with diarrhea in general and chronic diarrhea in particular in a cross-sectional study of HIV+ patients. The association between *C. canis*, *C. felis*, and subtype family Id of *C. hominis* infections and diarrhea were significant. On the other hand, infections with the *C. parvum* was more likely associated with chronic diarrhea and vomiting (Cama *et al.*, 2007).

Otherwise, associations between *C. hominis* subtype families and clinical presentations were very different. Nonetheless, these associations were seen in a longitudinal cohort study of cryptosporidiosis in children in the same area (Cama *et al.*, 2008). All subtype families were associated with diarrhea but the Ib subtype family appeared to be much more virulent, it being significantly associated with diarrhea, nausea, vomiting and general malaise during the first episode of infection and all infections. However, the Ia subtype family was also very pathogenic in children, being significantly associated with diarrhea, nausea and vomiting during the first episode of cryptosporidiosis (Cama *et al.*, 2008). These results demonstrated that *C. hominis* GP60 subtype families are linked to different clinical manifestations, children and HIV+ adults may have different clinical responses to infections with different *C. hominis* subtype families.

1.3.11 Human infections with other *Cryptosporidium* spp.

In the past, *C. parvum* was considered the only *Cryptosporidium* species to infect humans. Genotyping tools based on DNA sequences of antigen and house-keeping genes identified genotype 1 (the human genotype) and genotype 2 (the bovine genotype) and finally became *C. hominis* and *C. parvum*, both of which are infectious for immunocompetent and immunocompromised persons. Nowadays, in most countries, *C. hominis* and *C. parvum* are responsible for greater than 90% of human cases of cryptosporidiosis, with the balance attributable to *C. meleagridis*, *C. canis* and *C. felis*. However, developing countries have a high prevalence of these unusual species (Xiao and Feng, 2008). In Bangkok, Thailand and Lima, Peru, *C. parvum* is as prevalent in humans as *C. meleagridis*, responsible for 10-20% of human cryptosporidiosis cases (Cama *et al.*, 2008; Cama *et al.*, 2007; Gatei *et al.*, 2002). Most human infections with the *Cryptosporidium* cervine genotype were reported in industrialized nations whereas in developing countries, human *C. canis* infections were reported (Xiao and Feng, 2008).

In developing countries, some human infections with *C. parvum* and the unusual species may be the result of anthroponotic transmission. Using specific genotyping tools in Lima, Peru concurrent presence of *C. parvum* and *C. hominis* IIc subtype family was demonstrated in 6 of 21 HIV+ patients. Otherwise, Cama *et al.* (2006) reported that some of the *C. canis* and *C. felis* infections in humans were probably transmitted through the anthroponotic rather than zoonotic pathway. Peruvian children studied with multiple cryptosporidiosis episodes found that sequential infections

with heterologous *Cryptosporidium* species were more common than sequential infections with homologous *Cryptosporidium* species (Cama *et al.*, 2008). Some children had the same species (*C. hominis*) in sequential infections. In Lima, Peru, HIV+ patients infected with *C. parvum*, *C. canis* and *C. felis* were more likely to have chronic diarrhea; those infected with *C. canis* and *C. felis* were more likely to have diarrhoea in general and those with only *C. parvum* were also likely to have vomiting. On the one hand, infections with *C. meleagridis* were more likely asymptomatic (Cama *et al.*, 2007). However, children infected with *C. parvum*, *C. meleagridis*, *C. canis* and *C. felis* were less likely to have non-diarrheal symptoms such as general malaise, nausea and vomiting; only *C. hominis* infections were significantly associated with these clinical symptoms (Cama *et al.*, 2008).

1.3.12 *Cryptosporidium* species and genotypes in humans in developing countries

The PCR techniques used do not amplify DNA from some more genetically different *Cryptosporidium* species. Later, the use of small subunit (SSU) rRNA-based genotyping tools demonstrated the presence of *C. canis*, *C. felis*, and *C. meleagridis* in AIDS patients in many countries such as the United States, Switzerland, and Kenya. Moreover, *C. hominis* and *C. parvum* are more frequently found (Morgan *et al.*, 2000a; Morgan *et al.*, 1999; Morgan *et al.*, 2000b; Pieniazek *et al.*, 1999). This investigation has been supported by data from Thailand and Peru (Cacciò *et al.*, 2002; Gatei *et al.*, 2002; Tiangtip & Jongwutiwes, 2002). Even immunocompetent persons can be infected with zoonotic species other than *C. parvum*. Researchers found *C. meleagridis*, *C. felis*, and *C. canis* specimens in the United Kingdom (Leoni *et al.*, 2006b; Nichols *et al.*, 2006). This demonstrated that most of the infected persons were not immunocompromised. Other cases of *C. meleagridis* infection have been characterised in immunocompetent hosts in developing countries (Tables 1.1 and Table 1.2).

In Peru, there was no significant difference between children and HIV-1 infected adults in the distribution of *C. hominis*, *C. parvum*, *C. meleagridis*, *C. felis*, and *C. canis* (Cama *et al.*, 2003). Otherwise, *Cryptosporidium* muris-like oocysts were found in two healthy Indonesian girls (Katsumata *et al.*, 2000). However, the sequence was more similar to *C. andersoni* than to *C. muris*. *C. muris* infections in AIDS patients in Kenya and Peru was confirmed both by PCR restriction

fragment length polymorphism (RFLP) and sequencing of the SSU rRNA gene (Gatei *et al.*, 2006; Palmer *et al.*, 2003), and also *C. muris* infection was found in a human in India (Muthusamy *et al.*, 2006). More human cases have been associated with the *Cryptosporidium* cervine genotype (Leoni *et al.*, 2006b; Nichols *et al.*, 2006; Soba *et al.*, 2006; Trotz-Williams *et al.*, 2006). Other *Cryptosporidium* species found in humans include *C. suis* and *C. suis*-like parasite (Leoni *et al.*, 2006b; Nichols *et al.*, 2006; Ong *et al.*, 2002), *C. andersoni*-like parasite (Leoni *et al.*, 2006a; Morse *et al.*, 2007), the chipmunk genotype I(W17) and the skunk genotype (Feltus *et al.*, 2006; Nichols *et al.*, 2006). The *C. hominis* monkey genotype has also been found in humans (Mallon *et al.*, 2003). Some unusual *Cryptosporidium* species may have a wide host range and when the environment changes, they may emerge as important human pathogens. Otherwise, the avian pathogen *C. meleagridis* is increasingly recognised as an important human pathogen, *C. meleagridis* is responsible for 10-20% of human cryptosporidiosis cases in Lima, Peru, and Bangkok, Thailand (Cama *et al.*, 2003; Gatei *et al.*, 2002; Xiao *et al.*, 2001a).

1.3.13 Zoonotic cryptosporidiosis in developing countries

The distribution of *Cryptosporidium* spp. in humans in developing countries is very different from that in most industrialized nations. Many studies concluded that a dominance of *C. hominis* in humans in developing countries was responsible for 70-90% of infections (Bushen *et al.*, 2007; Cama *et al.*, 2003; Das *et al.*, 2006; Gatei *et al.*, 2007; Gatei *et al.*, 2006; Muthusamy *et al.*, 2006; Peng *et al.*, 2003; Tiangtip & Jongwutiwes, 2002; Tumwine *et al.*, 2003). In contrast, the disease burden attributable to *C. parvum* is much lower. This shows that zoonotic infection is much less common in developing countries than in industrialized countries. In fact, most human *C. canis* infections have been reported in persons in developing countries. In Peru and Thailand, *C. meleagridis*, *C. canis*, and *C. felis* are responsible for 15-20% of *Cryptosporidium* infections in AIDS patients and children (Table 1.2 and Table 1.3). In developing countries, the anthroponotic IIc subtype family is responsible for most human *C. parvum* infections (Akiyoshi *et al.*, 2006; Peng *et al.*, 2003; Xiao & Ryan, 2004). In some regions such as Lima, Peru, the IIc subtype family is the only *C. parvum* in humans, whereas in other developing countries such as Malawi and Kenya, IIe *C. parvum* subtype family is also seen in humans (Cama *et al.*, 2007; Xiao *et al.*, 2004; Xiao & Ryan, 2004). In Uganda, IIc subtypes are the dominant *C. parvum* in children (Akiyoshi *et al.*, 2006).

A recent study in Malawi has shown a higher *C. parvum* infection rate in rural areas than in urban areas (Morse *et al.*, 2007). Whether *C. meleagridis*, *C. canis*, *C. felis*, and *C. muris* are transmitted in developing countries by the zoonotic pathway remains to be decided. Using *C. hominis* and *C. parvum*-specific genotyping tools, the analysis of *C. canis* and *C. felis*-infected specimens from HIV1 persons in Lima, Peru, revealed the concurrent presence of the *C. hominis* and *C. parvum* IIc subtype family in 6 of 21 patients, indicating that infection with mixed *Cryptosporidium* spp. is more prevalent than believed previously. The concurrent presence of the human-specific *C. hominis* and *C. parvum* also suggests that many of the *C. canis* and *C. felis* infections in humans were transmitted through the anthroponotic rather than the zoonotic pathway (Cama *et al.*, 2006). There are no multilocus subtyping studies to determine whether there is any host segregation in *C. canis* or *C. felis*, although an earlier study of a small number of human and bird specimens failed to show this in *C. meleagridis* (Glaberman *et al.*, 2002).

1.3.14 Molecular epidemiology and public health implications

Various molecular diagnostic tools have been used to differentiate *Cryptosporidium* species/genotypes and *C. parvum* and *C. hominis* subtypes.

1.3.14.1 Small subunit (SSU) rRNA

These tools are generally used in genotyping *Cryptosporidium* in humans, animals and water samples. For genotyping, a PCR-RFLP tool and uses *SspI* and *VspI* restrictions is commonly used (Xiao *et al.*, 2001a; Xiao *et al.*, 2001b). Detecting *Cryptosporidium* genotypes using SSU rRNA gene methodology has proved to be an effective method. This is because of the presence of semi-conserved and hyper-variable regions on the gene and its multi-copying attributes which can make possible the creation of primers distinct to the genus.

1.3.14.2 Oocyst wall protein (COWP) gene

This gene has been used for the characterisations of *Cryptosporidium* spp., however, the use of this gene in *Cryptosporidium* genotyping has decreased recently. Generally, PCR tools based on COWP gene only amplify DNA of *C. parvum*, *C. hominis*, *C. meleagridis*, and species/genotypes closely related to *C. parvum*. Otherwise, COWP-based tools have limited usefulness in genotyping

Cryptosporidium spp. of animals because this gene has a low specificity. Subtyping tools are widely used for the study of the transmission of *C. hominis* in humans and *C. parvum* in humans and ruminants.

1.3.14.3 60 kDa glycoprotein (GP60 or GP40/15)

The GP60 gene has tandem repeats of the serine-coding trinucleotide TCA, TCG or TCT at the 50 (GP40) end of the gene which is similar to a microsatellite sequence. Furthermore, characterisation of *C. parvum* and *C. hominis* each to several subtype families, there are extensive sequence differences in the non-repeat regions with variations in the number of trinucleotide repeats. Within each subtype family, subtypes differ from each other mostly in the number of trinucleotide repeats (TCA, TCG or TCT microsatellite). Sequence analysis of the hypervariable GP60 gene, is extensively used in *Cryptosporidium* subtyping because of its sequence heterogeneity and relevance to parasite biology. Moreover, it is the most single polymorphic marker characterised in the *Cryptosporidium* genome (Leoni *et al.*, 2007; Wielinga *et al.*, 2008). In addition, it is different from other subtyping targets such as internal transcribed spacer-2, double stranded RNA, microsatellites and mini-satellites. Otherwise, it is one of the dominant targets for neutralising antibody responses in humans (O'Connor *et al.*, 2007). So, it is possible to link biologic characteristics of the parasites and clinical presentations with the subtype family identity. Some of the *C. parvum* subtype families such as IIa and IIc are found in both humans and ruminants, responsible for zoonotic cryptosporidiosis. There are also significant differences in clinical presentations and virulence among some common *C. hominis* subtype families in cryptosporidiosis-endemic areas (Cama *et al.*, 2008).

1.3.14.4 The identification of microsatellite and minisatellite sequences

Whole genome sequencing of *Cryptosporidium* was recently used to identify microsatellite and minisatellite sequences in *C. parvum* and *C. hominis* genomes, as well as other highly polymorphic targets of this parasite. They are frequently used to increase the subtyping resolution in multilocus analysis.

1.3.14.5 Multilocus Sequence Typing (MLST)

The second typing technique is multilocus sequence typing (MLST). The principle of this method, MLST, is used to accurately characterise strains of microbial species including bacteria. This characterisation uses the unique allelic profiles in a set of housekeeping genes. The technique requires PCR amplification followed by DNA sequencing. The allelic sequences of the organism are checked against all the previously recorded sequences on an internet database (multilocus sequence analysis). MLST allows the detection of length polymorphism and nucleotide substitution in microsatellite and mini-satellite markers and the inclusion of other markers with only single nucleotide polymorphisms (SNPs). However, MLST depends on the detection of genetic heterogeneity by DNA sequencing of the amplified PCR products (Cama *et al.*, 2006; Gatei *et al.*, 2008; Gatei *et al.*, 2007). Moreover, MLT and MLST tools are widely used in geographic tracking and genetic population studies of *C. parvum* and *C. hominis* (Gatei *et al.*, 2008; Morrison *et al.*, 2008; Tanriverdi *et al.*, 2008). Nonetheless, all extensively specific tools which are used in genotyping or subtyping *Cryptosporidium*, have the problem of identifying only the dominant genotype in the specimen. Otherwise, all narrowly specific genotyping and subtyping tools detect only *C. parvum*, *C. hominis* and species or genotypes related to them. Therefore, the type of molecular diagnostic tools used and their order of usage would vastly affect the results of molecular epidemiologic investigations of cryptosporidiosis transmission (Savioli *et al.*, 2006; Smith *et al.*, 2006). The use of a SSU rRNA-based PCR-RFLP genotyping tools resulted in the identification of *C. canis* and *C. felis* in multiple specimens from 21 HIV+ adults in Peru (Cama *et al.*, 2006).

The use of molecular methods has had a huge impact on elucidating the nature of variation in *Cryptosporidium*, especially because they enable the direct recognition of cysts or oocysts recovered from fecal or environmental samples (Robertson *et al.*, 2006; Savioli *et al.*, 2006; Smith *et al.*, 2006). To detect species, analysis of highly or moderately conserved coding regions (e.g. 18S rDNA, structural and housekeeping genes) is required, whereas investigations into the transmission of genotypes and subtypes, identifying sources of infection and risk factors, requires more discriminatory fingerprinting techniques, such as those based on mini-satellites (MN) and micro-satellites (MS), which can identify individual isolates or clonal lineages.

1.4 Detection of *Giardia duodenalis* and *Cryptosporidium* spp. in clinical samples

1.4.1 Microscopy

Light microscopy is routinely used for the detection of *Giardia duodenalis* in both humans and animals. Stool samples (the direct faecal smear) are normally used for checking the cysts and trophozoites of *G. duodenalis* (Hill, 2001). The examination of the faecal samples is usually performed at least three times from the same person. This is because the cysts are frequently excreted at intervals (sporadically) (Manser *et al.*, 2014). Occasionally, the trophozoites of *G. duodenalis* can be found in the duodenal aspirates of patients (Roberts & Zeibig, 2013). Sometimes the floatation method (concentrated zinc sulphate) and Formol-Ether (Ethyl Acetate) concentration method can be conducted for diagnosis of this parasite in case the direct faecal smear does not give a satisfactory result (Barr, 1998). Another method which is more sensitive than the microscopic technique is direct fluorescence assay (DFA), which may be useful for samples which have a small number of cysts.

Cryptosporidium spp. require special staining methods to enable their detection by microscopy as they are very small and difficult to distinguish. There are some staining techniques that are popular to use such as acid fast stain (AFS) and the Modified Ziehl Neelsen (MZN) staining technique. This technique is specific (Behr *et al.*, 1997; Feng *et al.*, 2007c) but it is not very sensitive (Verweij *et al.*, 2003).

1.4.2 Immunological methods

Other sensitive detection tools such as enzyme-linked immunosorbent assay (ELISA) (Barr *et al.*, 1998) with monoclonal or polyclonal antibody, Enzyme immune assays (EIA) and non-enzymatic immunochromatographic assays are used to diagnose parasite antigens from stool samples (Aldeen *et al.*, 1998; Mank *et al.*, 1997). These techniques are recommended for laboratory diagnosis for screening a larger number of faecal specimens and they are generally more sensitive than conventional light microscopy.

1.4.3 Molecular methods

Polymersase Chain Reaction (PCR) is a very popular technique for detection and identification of parasites as it is much more sensitive and specific than the methods outlined above (Thompson *et al.*, 2008). In the past decade, Real-time PCR assays have been developed to amplify the DNA of parasites from faeces (Almeida *et al.*, 2010; Amar *et al.*, 2002; Calderaro *et al.*, 2010; Manser *et al.*, 2014; Verweij *et al.*, 2003). This assay provides results that have a significantly higher sensitivity than microscopic methods. The sensitivity of the test by PCR was 100% whereas it was 86.7% by the microscopic method (Calderaro *et al.*, 2010). Recently, multiplex PCR assays have been developed which were able to detect a variety of micro-organisms (parasites, protozoa, bacteria and viruses) and a large number of faecal specimens (high throughput) can be performed at the same time (Verweij, 2014).

Previously, it was not possible to distinguish unidentified differences in disease, symptomatology, zoonotic potential, risk factors and environmental contamination for *Giardia* and *Cryptosporidium*. Nowadays, we can recognise species, genotype and also subtype easily by using molecular tools. Moreover, these tools can be used for both the identification of species and genotypes in the faeces of infected hosts and their recognition in environmental samples. Molecular biology has provided insights for characterising and understanding the epidemiology of *Giardia* and *Cryptosporidium* (Harp, 2003; Mak, 2004).

Using molecular diagnostic tools in the taxonomy of both genera has led to increased recognition of the diversity of the species infecting humans.

1.5 Current status of *Giardia duodenalis* and *Cryptosporidium* spp. infection in Africa

1.5.1 *Giardia duodenalis* (Table 1.3)

583 Rwandan children under the age of five were chosen to be studied to discover how common the rate of *G. duodenalis* infection was. The subjects were chosen from areas that were endemic for its presence. Of those children who presented with diarrhea and other associated symptoms such as stomach pain and nausea, faecal specimens were collected. Microscopy found a rate of 19.8% positive for *G. duodenalis*, whereas PCR found approximately 60% positivity (Ignatius *et al.*, 2012).

Moreover, the most predominant was *G. duodenalis* assemblage B infection. However, the prevalence of *G. duodenalis* in this area is probably highly underestimated due to the use of the low sensitivity microscopic method, especially when stool samples were only investigated once (Ignatius *et al.*, 2012). In a study in Ghana, samples collected from the gastrointestinal and urinary tracts of HIV seropositive patients were assessed for the presence of *G. duodenalis* and *Cryptosporidium* spp. Microscopically, 19% (95/500) of all patients were infected with *G. duodenalis* while *Cryptosporidium* spp. were detected in 14% (70/500) using the Modified Ziehl Neelsen (MZN) staining technique. It was concluded that *G. duodenalis* and *Cryptosporidium* were the most common diarrhoeal diseases. However, other intestinal parasites were also detected but in very low numbers. Based on the age and sex of the patients, these other parasites made no statistically significant difference (Boaitey *et al.*, 2012).

In Zambia, a study by Siwila *et al.* (2011) on seasonal variation and incidence of *Cryptosporidium* and *G. duodenalis* in children without diarrhoea from four pre-schools in Kafue was carried out. Using Direct Immunofluorescence Assay, *Cryptosporidium* spp. was found in 30.7% (241/786) of children whereas *G. duodenalis* was detected in 29.0% (228/786). *Cryptosporidium* spp. and *G. duodenalis* infections were found to be higher in the wet season than in the dry season. Cryptosporidiosis was significantly correlated with diarrhoea whereas giardiasis was not associated with this disease (Siwila *et al.*, 2011). In Uganda, the prevalence of intestinal protozoa was investigated from faecal specimens of people living in six Ugandan communities by using a rapid antigen test; the prevalence of *Giardia* was 12% (McElligott *et al.*, 2013).

In South Ethiopia, in a study on the prevalence of intestinal parasitic infections by using direct faecal smear and formalin-ether concentration techniques, *G. duodenalis* was detected in 10.6% (91/858). Sex was not correlated with parasitic infections but there was a high prevalence of intestinal parasitic infections in the reproductive age group (Wegayehu *et al.*, 2013b). In Ethiopia, there was a high prevalence of intestinal parasitic infections among HIV/AIDS patients. In particular, a high infection rate of *Cryptosporidium* spp. and *G. duodenalis* was found in HIV/AIDS patients with CD4⁺ T cell counts of <200 cells/mL. Generally, HIV patients not receiving antiretroviral treatment (ART) were infected with at least one intestinal parasite species (Adamu *et al.*, 2013b).

In Tanzania, identification and genetic characterisation of *G. duodenalis* in 45 human faecal specimens from children in primary schools was performed by using SSU-rDNA and *gdh* genes. *G. duodenalis* assemblage B (48.9%) was of a higher prevalence than *G. duodenalis* assemblage A (6.66%) (Di Cristanziano *et al.*, 2014). However, this was a very small sample size, more samples should have been collected than this so that the results and the ratio of both assemblages would be more reliable. In Uganda, *G. duodenalis* was identified in 20.1% (86/427) of healthy children as part of a previous *Helicobacter pylori* survey. Door to door visits were made in urban Kampala, Uganda and information was obtained from children aged between 0 and 12 years old. The highest rates were found in infants aged 15 years old. There was no significant association in non-symptomatic Ugandan children among *Giardia* infection and gender, types of toilet and housing or types of drinking water source. 45 *G. duodenalis* positive samples were characterised using multi-locus genotyping analysis (*bg*, *gdh* and *tpi*), 11.1% (5/45) with *G. duodenalis* assemblage AII, 55.6% (25/45) with *G. duodenalis* assemblage B, and 8.9% (4/45) with mixed assemblage infections (A/B). The *G. duodenalis* assemblage B is predominant in Uganda (Ankarklev *et al.*, 2012). Nevertheless, the sample size of this study is small.

Table 1.3 Summary of studies of *Giardia duodenalis* infection in Africa

Country	Area of Study (Province)	Subjects	Prevalence (%) reported	Method	Reference
Rwanda	Rwandan	children <5 years of age	19.8% 60%	Microscopy PCR	Ignatius <i>et al.</i> (2012)
Ghana	Komfo Anokye Teaching Hospital	HIV seropositive patients	19% (95/500)	Microscopy	Boaitey <i>et al.</i> (2012)
Zambia	Kafue	4 pre-schools children	<i>G. duodenalis</i> 29.0% (228/786).	Immunofluorescence Assay	Siwila <i>et al.</i> (2011)
Uganda	six Ugandan communities	-	12%	a rapid antigen test	McElligott <i>et al.</i> (2013)

South Ethiopia	-	HIV/AIDS patients	10.6% (91/858)	direct fecal smear and formalin-ether concentration techniques	Wegayehu <i>et al.</i> (2013b)
Ethiopia	-	HIV patients	a high infection rate of <i>Cryptosporidium</i> spp. and <i>G. duodenalis</i>	-	Adamu <i>et al.</i> (2013b)
Tanzania	-	children in primary schools	- assemblage B (48.88%) - assemblage A (6.66%)	PCR, SSU-rDNA and <i>gdh</i> genes	Di Cristanziano <i>et al.</i> (2014)
Uganda	door-to-door visits in urban Kampala	healthy children (aged between 0-12 years)	- 20.1% (86/427) - 11.11% (5/45) with <i>G. duodenalis</i> assemblage AII - 55.55% (25/45) assemblage B - 8.88% (4/45) with mixed assemblage infections (A/B)	multi-locus genotyping analysis (<i>bg</i> , <i>gdh</i> and <i>tpi</i>)	Ankarklev <i>et al.</i> (2012)

1.5.2 *Cryptosporidium* spp.

In Africa, *Cryptosporidium* spp. infection has been correlated with 3.8–26% of diarrhoea in young children (Amadi *et al.*, 2001; Mwachari *et al.*, 1998) and it has been detected in about 47% of African countries. In the past, there was a survey of 186 paediatric patients with gastro-enteritis who attended an out-patient clinic in Malawi, the prevalence of *Cryptosporidium* was 42%. (Pavone *et al.*, 1990). *Cryptosporidium* spp. infection was detected in approximately 10% of infants under 5 years of age with diarrhoeal illness. This was, however, higher in immunocompromised children where an infection rate of 21% was detected (Pavone *et al.*, 1990). Cranendonk *et al.* (2003) investigated *C. parvum* infection in diarrhoea in-patients at Queen Elizabeth Central Hospital (QECH) in Blantyre, Malawi. *C. parvum* were examined by an immuno-fluorescent assay. The prevalence of *C. parvum* was 11% (13/121) (Cranendonk *et al.*, 2003). A study on intestinal parasitic infections in HIV positive and HIV negative patients in Cameroon found that the overall prevalence of intestinal parasites was 14.6%. HIV/AIDS patients were infected with these parasites 59.5% (25/42) whereas the HIV negative patients were infected 9.3% (33/354). In the HIV infected

patients, 19% were infected with *C. parvum* and 2.4 with *G. duodenalis*. Otherwise, in the non-HIV infected patients, 2.5% with *C. parvum*, and 0.25% with *G. duodenalis* were found. *C. parvum* was the most predominant in HIV/AIDS patients compared to other parasites in west Cameroon (Nkenfou *et al.*, 2013). In Ethiopia, Zeynudin *et al.* (2013) investigated the prevalence of opportunistic intestinal parasitic infection in HIV patients, 91 samples were detected by the use of MZN staining. The prevalence of *C. parvum* was 13.3%. The prevalence of opportunistic parasites (*I. belli*, *C. parvum*, *Cyclospora cayetanensis* and Microsporidia species) was 15.4%. (Zeynudin *et al.*, 2013). In Egypt, Abd *et al.* (2012) detected and characterised *Cryptosporidium* spp. in human patients in Cairo. A total of 391 human faecal samples were screened by immunochromatographic detection kit; Stick *Crypto-Giardia* (Operon) “the Stick *Crypto-Giardia*; Operon”. *Cryptosporidium* spp. were found in 23 of 391 samples (5.9%). The oocyst wall protein (COWP) gene and RFLP analysis were used for genotyping, 78.4% (18/23) were positive. Only 15 samples were characterized, nine samples (60%) were *C. hominis*, three samples were *C. parvum* (20%) and three samples were mixed infection (*C. hominis*/*C. parvum*, 20%) (Abd *et al.*, 2012). Abdel-Hafeez *et al.* (2012) studied opportunistic intestinal protozoa from immunosuppressed infants who presented themselves to the clinic and were admitted to the clinics of the Paediatric Department Minia University Hospital, Egypt. The patients were divided into two groups; immunosuppressed (200 samples) and immunocompetent group (250 samples). Direct smear method and formol-ether concentration technique were performed. MZN and Giemsa were used for staining. Intestinal protozoa infection was found in 188 of 200 samples (94%) among immunosuppressed paediatric, whilst 60% (150/250) were infected with intestinal protozoa in immunocompetent infants. *C. parvum* had the highest prevalence in immunosuppressed group: 60.2% were infected with *C. parvum*. In contrast, in immunocompetent paediatrics, 24.6% were infected with *E. histolytica* and 17.6% with *G. duodenalis*. It can be concluded that *E. histolytica* and *G. duodenalis* were common in immunocompetent infants whereas *C. parvum* was the most common in immunosuppressed children (Abdel-Hafeez *et al.*, 2012). Adamu *et al.* (2010) collected 1,034 stool specimens from patients with diarrhoea in nine different areas in Ethiopia for identification and characterization of *Cryptosporidium* spp. Oocysts of *Cryptosporidium* were detected in 7.6% (79/1034) of patients by MZN staining. DNA extraction was performed on 79 positive samples, which were then amplified by PCR using COWP, SSU-

rRNA and GP60 gene. *Cryptosporidium* spp. were successfully amplified and sequenced in 52% (41/79). Of these positive for *Cryptosporidium* spp., 39 samples were characterized as *C. parvum* (95.1%), 2.4% (1/41) with *C. hominis* and 2.4% (1/41) with mixed infection. These results demonstrated that *C. parvum* is the major cause of human cryptosporidiosis in Ethiopia and indicate a zoonotic transmission of the disease. This is because three subgenotypes of *C. parvum* were found by sequencing of the GP60 gene, all subtypes of these isolates belonged to the zoonotic allele families IIa and one subtype belonged to *C. hominis* (Ib). It is different from previous research in various countries (Adamu *et al.*, 2010). Mbae *et al.* (2013) studied intestinal parasitic infections in young patients under 5 years of age with diarrhoea that were admitted and presented to Mbagathi District hospital in Nairobi, Kenya. Direct wet mount, formal-ether sedimentation technique and the MZN staining technique were used in this study. The overall prevalence was 25.6% (541/2112). Of those positive for parasitic infections, 36.7% were infected with *E. histolytica*, 30.5% with *Cryptosporidium* spp. and 16% with *G. duodenalis*. Intestinal parasite infection in infants who presented themselves to the clinic at the hospital had higher prevalence (27.4%, 432/1577) than those in inpatients (20.1%, 109/535). It showed that *E. histolytica* and *Cryptosporidium* spp. had a high prevalence in this country (Mbae *et al.*, 2013).

In South Africa, Samra *et al.* (2012) studied the genetic characterization of *Cryptosporidium* spp. in diarrhoeic children. 442 faecal specimens from paediatrics under 5 years of age were investigated by using MZN acid-fast staining. *Cryptosporidium* spp. was found in 12.2% (54/442). Of these, 25 samples were successfully amplified by PCR, and then were genotyped by RFLP and DNA sequence analyses of the 18S rRNA gene. 76% were characterised as *C. hominis*. Five *C. hominis* subtype families were Ia, Ib, Id, Ie and If. Of these, 20% were identified as *C. parvum* with three subtype families: IIc, IId and IIb. From the latter subtypes, the most common was subtype family IId, while one subtype family was *C. meleagridis* (IIId) (Samra *et al.*, 2012). In the rural region of Northern Tunisia, 403 stool specimens from infants were determined by MZN staining and PCR-RFLP was performed by using an 18S rRNA nested PCR. Seven human samples were detected by the MZN method. Four were identified as *C. parvum* and three were identified as *C. meleagridis*. The IIaA15G2R1 and the IIdA16G1 were characterised for intraspecific of *C. parvum* (Rahmouni *et al.*, 2014).

In Western Uganda, Salyer *et al.* (2012) investigated risk factors for Cryptosporidiosis and assessed cross-species transmission using a nested PCR (COWP gene). Faecal samples from human, primates and cattle were collected. A prevalence of *Cryptosporidium* in humans, in non-human primates and in livestock were 32.4%, 11.1%, and 2.2%, respectively. There was no *Cryptosporidium* and *G. duodenalis* in the same samples (no co-infection) (Salyer *et al.*, 2012). In Tanzania, Moyo *et al.* (2011) determined the age-specific aetiologic agents of diarrhoeal disease in paediatric under five years of age (in-patients) from four hospitals in Dar es Salaam, Tanzania. Intestinal protozoa were examined using an Enzyme Linked Immunosorbent Assay (ELISA). The prevalence of *C. parvum* was 18.9%, Acute watery diarrhoea was found higher in the dry season compared to the wet season in children less than five years of age in Dar es Salaam (Moyo *et al.*, 2011).

Table 1.4. Summary of studies of *Cryptosporidium* spp. infection in Africa

Country	Area of Study (Province)	Subjects	Prevalence (%) reported	Methods	References
Zambia	University Teaching Hospital in Lusaka	Patients with AIDS-related diarrhoea 200 children (94 boys and 106 girls, 6-24 months old)	26% (malnutrition and show high levels of mortality)	-	Amadi <i>et al.</i> (2001)
Kenya	Nairobi	diarrhoea in young children <5 years of age	17%	Ziehl-Neelsen staining and light microscopy	Mwachari <i>et al.</i> (1998)
Malawi	-	186 paediatric patients with gastro-enteritis (out-patient clinic)	42%	Microscopy	Pavone <i>et al.</i> (1990)
Malawi	-	infants under 5 years of age with diarrhoeal illness.	10%	-	Pavone <i>et al.</i> (1990)
Malawi	-	immunocompromised children	21%	-	Pavone <i>et al.</i> (1990)
Zambia	Kafue	4 pre-schools children	30.7% (241/786)	Immunofluorescence Assay	Siwila <i>et al.</i> (2011)

Ghana	-	HIV seropositive patients	14% (70/500)	Modified Ziehl Neelsen (MZN) staining technique.	Boaitey <i>et al.</i> (2012)
Malawi	Queen Elizabeth Central Hospital (QECH) in Blantyre,	diarrhoea in-patients	11% (13/121) was <i>C. parvum</i>	Immuno-fluorescent assay	Cranendonk <i>et al.</i> (2003)
Cameroon	-	HIV positive and HIV negative patients	-14.64% HIV/AIDS, 59.5% (25/42) HIV negative patients, 9.32% (33/354)	-	Nkenfou <i>et al.</i> (2013)
Cameroon	-	In the HIV infected patients, In the non-HIV infected patients,	19.04% with <i>C. parvum</i> and 2.38% with <i>G. duodenalis</i> 2.53% with <i>C. parvum</i> , and 0.25% with <i>G. duodenalis</i>	-	Nkenfou <i>et al.</i> (2013)
Ethiopia	-	HIV patients (91 samples)	13.3% (<i>C. parvum</i>)	MZN staining	Zeynudin <i>et al.</i> (2013)
Egypt	human patients in Cairo	391 human faecal samples	5.9% (23/391)	immunochromatographic detection kit; Stick <i>Cryptogiardia</i> (Operon)	Abd <i>et al.</i> (2012)
Egypt	human patients in Cairo	23 samples 15 samples were characterized	78.36 % (18/23) 60% (9/15) were <i>C. hominis</i> , 20% (3/15) were <i>C. parvum</i> , 20% (3/15) were mixed infection (<i>C. hominis/C. parvum</i>)	The oocyst wall protein (COWP) gene and RFLP analysis	Abd <i>et al.</i> (2012)
Egypt	the clinics of the Paediatric Department Minia University Hospital	immunosuppressed infants immunosuppressed (200 samples) and	-188 of 200 samples (94%) among immunosuppressed paediatric	Direct smear method and formol-ether concentration technique were performed. MZN and Giemsa	Abdel-Hafeez <i>et al.</i> (2012)

		immunocompetent group (250 samples)	- 60% (150/250) in immunocompetent infants		
Egypt	the clinics of the Paediatric Department Minia University Hospital	In immunocompetent paediatrics	(60.2%) was <i>C. parvum</i>	Direct smear method and formol-ether concentration technique were performed. MZN and Giemsa	Abdel-Hafeez <i>et al.</i> (2012)
Ethiopia	-	1,034 stool specimens from patients with diarrhoea in nine different areas	7.6% (79/1034)	MZN staining	Adamu <i>et al.</i> (2010)
Ethiopia	-	79 positive samples	52% (41/79) 39 samples were characterized as - <i>C. parvum</i> (95.1%) - <i>C. hominis</i> (2.4%, 1/41) - mixed infection (2.4%, 1/41)	PCR using COWP, SSU-rRNA and GP60 gene.	Adamu <i>et al.</i> (2010)
Kenya	Mbagathi District hospital in Nairobi	young patients under 5 years of age with diarrhoea	30.5%	Direct wet mount the MZN staining technique	Mbae <i>et al.</i> (2013)
Kenya	Mbagathi District hospital in Nairobi	young patients under 5 years of age with diarrhoea	infants 27.4% (432/1577) inpatients 20.1% (109/535).	Direct wet mount the MZN staining technique	Mbae <i>et al.</i> (2013)
South Africa	-	diarrhoeic children. 442 faecal specimens, paediatrics under 5 years of age	12.2% (54/442)	MZN acid-fast staining	Samra <i>et al.</i> (2012)
South Africa	-	76% were characterised as <i>C. hominis</i> .	Five <i>C. hominis</i> subtype families were Ia, Ib, Id, Ie and If. Of these, 20% were identified as <i>C. parvum</i> (IIc, Iie and Iib) The most	PCR, RFLP and DNA sequence analyses of the 18S rRNA gene	Samra <i>et al.</i> (2012)

			common was subtype family IIe, one subtype family was <i>C. meleagridis</i> (IIIId)		
Tunisia	the rural region of Northern	403 stool specimens from infants	Seven human samples	MZN staining and PCR-RFLP	Rahmouni <i>et al.</i> (2014)
Tunisia	the rural region of Northern	Seven human samples	<i>C. parvum</i> (4 samples; IIaA15G2R1 & IIaA16G1 <i>C. meleagridis</i> (3 samples	18S rRNA nested PCR	Rahmouni <i>et al.</i> (2014)
Western Uganda	-	Faecal samples from humans, primates and cattle	Humans (32.4%), non-human primates (11.1%), livestock (2.2%)	nested PCR (COWP gene)	Salzer <i>et al.</i> (2012)
Tanzania	in-patients from four hospitals in Dar es Salaam	diarrhoeal disease in paediatric under five years of age	18.9%	Enzyme Linked Immunosorbent Assay (ELISA) and agglutination assay	Moyo <i>et al.</i> (2011)

1.6 Current status of *Giardia duodenalis* and *Cryptosporidium* spp. infection in Southeast Asia

1.6.1 *Giardia duodenalis* (Table 1.5)

Cambodia is recognised as one of South East Asia's poorest countries with poor health and significant ratios of malnutrition (Marmot & Friel, 2008). A large population live in poor hygienic conditions (UNICEF, 2009). Most of the households drink water obtained from rainwater (Chesnaye *et al.*, 2011). Therefore, children are often infected with intestinal parasites. The national Centre for Parasitology, Entomology and Malaria Control (NMC, MoH, Cambodia) investigated the prevalence of intestinal parasitic infection in school children (more than 6,600 students) and more than 50% of infection rates of soil-transmitted helminths (STH) were found as well as infection rates above 70% being detected in several areas (Sinoun, *et al.*, 2005). Previous studies found infection rates of STH ranged from 25 to 52 % (Lee *et al.*, 2004; Lee *et al.*, 2002).

Cambodia is also recognised as an endemic area for malaria and schistosomiasis (Urbani *et al.*, 2002) but there are few published papers that investigated the infection rates of intestinal parasites in Cambodia (Moore *et al.*, 2012; Moore *et al.*, 2015). There have been some reports outlining the intestinal parasitic infections of children and refugees in Cambodia (Gyorkos *et al.*, 1992; Lee *et al.*, 2002; Lurio *et al.*, 1991; Moore *et al.*, 2012; Moore *et al.*, 2015).

In a study on the prevalence of giardiasis in Thailand, both the paediatric population of orphanages have been studied (Janoff *et al.*, 1990; Mungthin *et al.*, 2001; Saksirisampant *et al.*, 2012; Saksirisampant *et al.*, 2003) and that of school children (Boontanom *et al.*, 2011; Ratanapo *et al.*, 2008; Saksirisampant *et al.*, 2012; Wongstitwilairoong *et al.*, 2007). The prevalence ranged from 5 to 37.7%. The greatest prevalence of *G. duodenalis* infection was at the orphanages (37.7%) (Saksirisampant *et al.*, 2003). In addition, school children less than 5 years of age had the biggest prevalence and this prevalence was slightly higher in males than females (Ratanapo *et al.*, 2008). A study on the prevalence of giardiasis in orphanages was carried out in Bangkok between 1999 and 2001. This study demonstrated that the prevalence of *G. duodenalis* was 12% in 1990 but had slightly decreased in 2001 (10.6%) (Mungthin *et al.*, 2003). Tungtrongchitr *et al.* (2010) collected 61 faecal specimens (with and without symptomatic diarrhoea) from Bangkok and/or in rural areas of Thailand. In the genotyping of *G. duodenalis* in both areas, the assemblage B was found 51% (31/61), mixed infection of both assemblages was detected in 41% (25/61) and the assemblage A was detected in 8% (5/61) using the *tpi* gene. The prevalence of assemblage B and assemblage A and B combined were significantly higher than assemblage A. By using PCR-RFLP analysis and the *bg* gene found 12% contained assemblage AI, 88% displayed assemblage AII. RFLP analysis and the *gdh* gene found 45% contained assemblage BIII, 54.5% displayed assemblage BIV. This study found that all symptomatic cases were infected with AI (100%) and BIII sub-assemblages (50% of patients). In addition, the majority of symptomatic cases in Bangkok were adults and elderly (80%) whereas the main incidences of symptomatic cases in the rural specimens were children (85%) (Tungtrongchitr *et al.*, 2010).

In a study in India, Mukhopadhyay *et al.* (2007) collected faecal specimens from 253 children who had persistent diarrhoea, 155 children who had acute diarrhoea (disease control group) and from

100 children who had no diarrhoea (normal control group). Of those positive for parasitic infections in children who had persistent diarrhoea, *G. duodenalis* had the highest prevalence (67.7%, 171/253). This study found associations between HIV infection, severe malnutrition and protozoal agents (*C. cayetanensis* and *Cryptosporidium* spp.) (Mukhopadhyay *et al.*, 2007).

A study by Behera *et al.* (2008), one hundred faecal specimens were collected from adults (50 samples) and children (50 samples) who had malabsorption and another group with normal controls (50 samples) who had no diarrhea. Wet mount and Kinyoun's modified acid-fast staining (Behera *et al.*, 2008). The overall prevalence of intestinal parasitic infections was 80% (40/50) in adults and 82% (41/50) in children. Of those positive for parasitic infections in adults, 24% (12/50) were infected with *G. duodenalis*. The parasites diagnosed in children, 16% (8/50) were infected with *G. duodenalis* and 14% (7/50) with *Cryptosporidium* spp. Intestinal coccidian had a clear correlation to malabsorption syndrome (malnourished children) in this study (Behera *et al.*, 2008).

Another study in India, *Giardia* diarrhoea in a birth cohort of 452 children, in an urban slum in South India, was identified. Children with diarrhea from various infections (co-infections) tended to have a slightly longer duration of diarrhoea than children who only had diarrhea from *Giardia* infection. Analysis of PCR-RFLP at *tpi* gene, both in giardial diarrhoea (80%, 362/452) and asymptomatic giardiasis (94%, 425/452) was highly associated with Assemblage B. There were correlations between Assemblage A subgroup II alone or dual infections with both assemblage A and B having diarrhoea (Ajjampur *et al.*, 2009).

Moore *et al.* (2012) collected 16,372 faecal specimens from symptomatic paediatrics who attended at Angkor hospital for Children in Siem Reap, Cambodia. The samples were examined by microscopy. The overall prevalence of intestinal parasitic infections was 19.1% (3,121/16,372). Of those positive for parasitic infections, 8% (1,309/16,372) were infected with *G. duodenalis* (Moore *et al.*, 2012). Moore *et al.* (2015) examined intestinal parasite infections in symptomatic paediatrics who attended at Angkor hospital in Siem Reap, Cambodia. Faecal samples were collected from April to June 2012 and the samples were investigated by using concentrated faecal sample and microscopic method. For hookworm and *S. stercoralis*, two culture techniques, the agar plate culture and the charcoal coproculture method, were used. The overall prevalence was 39.3% (340/865). Of those positive for parasitic infections, 14.4% (124/865) were infected with hookworm,

11.6% (100/865) with *S. stercoralis* and 11.2% (97/865) with *G. duodenalis*. Children infected with *G. duodenalis* were significantly associated with the age 1-5 years whereas children infected with hookworm and *S. stercoralis* were associated with the age range of 5 years and older. Otherwise, hookworm, *S. stercoralis*, *G. duodenalis* were associated with children who lived outside of Siam Reap.

Table 1.5 Current status of *Giardia duodenalis* infection in Asia

Country	Area of Study (Province)	Subjects	Prevalence (%) reported	Method	Reference
Thailand	-	The paediatric population of orphanages (School Children)	5 to 37.7%.	Microscopy PCR	Boontanom <i>et al.</i> (2011); Ratanapo <i>et al.</i> (2008) Saksirisampant <i>et al.</i> (2012) Wongstitwilairoong <i>et al.</i> (2007)
Thailand	-	The orphanages	37%	-	Saksirisampant <i>et al.</i> (2003)
Thailand	Bangkok between 1990-2001	orphanages	12% in 1990 10.6% in 2001	-	Mungthin <i>et al.</i> (2003)
Thailand	Bangkok and/or in rural areas	61 faecal specimens (with and without symptomatic diarrhoea)	- The assemblage B was 51% (31/61), - mixed infection of both assemblages was detected 41% (25/61) -The assemblage A was 8% (5/61)	Genotyping of <i>G. duodenalis</i> and using the <i>tpi</i> gene.	Tungtrongchitr <i>et al.</i> (2010)
Thailand	Bangkok and/or in rural areas	-	- 12% (assemblage AI) - 88% displayed assemblage AII.	PCR-RFLP analysis and the <i>bg</i> gene	Tungtrongchitr <i>et al.</i> (2010)

Thailand	Bangkok and/or in rural areas	-	- 45% contained assemblage BIII - 54.5% displayed assemblage BIV	RFLP analysis and the <i>gdh</i> gene	Tungtrongchitr <i>et al.</i> (2010)
India	-	253 children (persistent diarrhoea)	67.7% (171/253)	-	Mukhopadhyay <i>et al.</i> (2007)
India	-	malabsorption - adults (50 samples) and - children (50 samples) normal controls (50 samples)	- 24% (12/50) was adults - 16% (8/50) children	Wet mount and Kinyoun's modified acid- fast staining	Behera <i>et al.</i> (2008)
India	in an urban slum in South India,	diarrhoea in a birth cohort of 452 children	- 80% (362/452) was giardial diarrhoea - 94% (425/452) was asymptomatic giardiasis	PCR-RFLP at <i>tpi</i> gene	Ajjampur <i>et al.</i> (2009)
Cambodia	Angkor hospital for Children in Siem Reap	16,372 faecal specimens from symptomatic paediatrics	8% (1,309/16,372)	microscopy	Moore <i>et al.</i> (2012)
Cambodia	Angkor hospital for Children in Siem Reap	symptomatic paediatrics	11.2% (97/865)	microscopic method	Moore <i>et al.</i> (2015)

1.6.2 *Cryptosporidium* spp. (Table 1.6)

In Nepal, *Cryptosporidium* was found in 5.6% (9/160) of children (5 years of age) with acute diarrhoea (Shariff *et al.*, 2002). In a study in India, Das *et al.* (2006) examined 1,338 human stool samples with diarrhoea and without diarrhoea. By microscopic method, oocysts of *Cryptosporidium* were detected in 3% (40/1,338). Of these positive for *Cryptosporidium* spp., 4.6% were found in humans who had diarrhoea and 1.2% were found in humans who had no diarrhea. PCR-RFLP of the 18S rRNA gene and sequencing were used for genotyping. *C. hominis*, *C. parvum* and *C. meleagridis* were characterised. *Cryptosporidium* spp. were successfully amplified and sequenced in 40 samples. Of these positive for *Cryptosporidium* spp., 35 samples were characterised as *C. hominis* (93.5%), 10% (4/40) with *C. parvum* and 2.5 % (1/40) with *C. felis*. These results indicated that *C. hominis* is the major cause of human cryptosporidiosis in India (Das *et al.*, 2006). One year later, Ajampur *et al.* (2007) studied very similarly to Das *et al.* (2006) but collected samples from Southern India. PCR-RFLP of the 18S rRNA gene was used for characterisation of *Cryptosporidium* spp. and Geographical Information Systems technology was used to provide spatial analysis of cases. Out of 53 positive samples; of these, 43 were positive for *C. hominis* (81%), 6 were positive for *C. parvum* (12 %), the rest of them (4) were positive for *C. felis* and *C. parvum* (mouse genotype) (7%). At the Cpgp 40/15 locus, 5 subgenotypes were characterised. Among *C. hominis* isolates, subgenotype Ia predominated and Ic was subgenotyped in all *C. parvum* isolates. There was association between *C. hominis* infection and a greater severity of diarrhoea. Sequencing of the Cpgp 40/15 alleles of *C. felis* and *C. parvum* (mouse genotype) showed similarities to subgenotype IIa and *C. meleagridis*. Space-time analysis recovered two clusters of infection due to *C. hominis* Ia, with a peak in February 2005. This study showed space-time clustering of a single sub genotype of *C. hominis* in a setting where cryptosporidiosis is endemic (Ajampur *et al.*, 2007). Gatei *et al.* (2007) examined 50 child patients (*Cryptosporidium* positive samples) in Kolkata, India. Multilocus Sequence Typing (MLST) was used for analysis. The results showed the presence of *C. hominis* (49/50), *C. meleagridis* (2/50), and *C. felis* (1/50) (there were mixed infections of *C. hominis* and *C. meleagridis* in 2 patients) (Gatei *et al.*, 2007). In another study in India, Muthusamy *et al.* (2006) identified *Cryptosporidium* spp. in 48 humans who were infected

with HIV. Multilocus genotyping was used for characterisation and *C. hominis*, *C. parvum*, *C. felis*, *C. mulis*, *C. meleagridis* were found.

Meamar *et al.* (2007) collected 24 stool specimens (15 human samples and 9 animal samples). Acid fast staining was used for detection and RFLP of the 18S rRNA gene and sequencing was used for genotyping (Xiao & Ryan, 2004). *C. hominis*, *C. parvum* and *C. meleagridis* were characterised. These results demonstrated that *C. parvum* is the major cause of human cryptosporidiosis in Iran and indicate a zoonotic transmission of the disease. Nonetheless, the sample size of this study is very small. It would be better to confirm the results by using more specimens (Meamar *et al.*, 2007). In Bangladesh, a study on genetic diversity of *Cryptosporidium* spp. in children was performed for the first time in the year 2011. Seven subtype families of *C. hominis* and *C. parvum* and 15 subtypes were found. The majority of specific families and subtypes was different from previous studies in various countries (Hira *et al.*, 2011). Ghaffari and Kalantari (2014) used a multilocus (various molecular techniques) for characterisation of *Cryptosporidium* spp. and to examine genetic variation of *Cryptosporidium* spp. isolated from Iran, Malawi, Nigeria, Vietnam and the United Kingdom. DNA samples were obtained from 106 positive specimens, which were then amplified by PCR, PCR-RFLP and DNA analysis of 18S rRNA, COWP, and GP60 gene. *C. hominis* and *C. parvum* were identified in the UK. Using sequence analysis of the GP60 gene, *C. hominis* (Ib, Ib3 & Id), *C. parvum* (IIa, IIc & IId) subtypes were characterised. The highest rate was *C. hominis* Ib (*C. hominis* Ib A10G2). From this study, *C. parvum* IId subgenotype was recorded for the first time. Three new subtypes of *C. parvum* IIa were identified in UK, a new subtype of *C. hominis* Id was found in Malawi.

Liu *et al.* (2014) collected 252 stool samples from children with diarrhea in a paediatric clinic (169 samples) and an intestinal clinic (83 samples) in China. The nested PCR was performed to characterise the parasites. The overall prevalence of infection was 30.2% (76/252). Of those positive for parasitic infections, 13.5% (34/252) were infected with *Cryptosporidium* spp. 13.5% (34/252) with *Enterocytozoon bieneusi* and 6.8% (17/252) with *G. duodenalis*. From sequence analysis, *C. andersoni* was characterised from all *Cryptosporidium* positive samples. The highest prevalence of *G. duodenalis* assemblage was assemblage C whereas only one sample was assemblage B. Co-infection between *Cryptosporidium* spp. and *E. bieneusi* was found in 8 patients whereas only

one sample had co-infection between *Cryptosporidium* spp. and *G. duodenalis*. This study found *Cryptosporidium* spp. and *E. bienersi* in winter more than in spring in China (Liu *et al.*, 2014).

The number of studies of *Giardia*, *Cryptosporidium* spp. and other parasites in humans in Southeast Asia is small (Shariff *et al.*, 2002; ; Das *et al.*, 2006; Ajjampur *et al.*, 2007; Gatei *et al.*, 2007; Meamar *et al.*, 2007; Nuchjangreed *et al.*, 2008; Hira *et al.*, 2011; Ghaffari & Kalantari, 2014; Liu *et al.*, 2014; Moore *et al.*, 2015). Furthermore, many research projects have been carried out on general information of intestinal parasites such as the incidence of infections and infection rates of *Giardia*, *Cryptosporidium* spp. and other protozoa parasites. Several studies were based on the microscopic method (Mak, 2004; Smith *et al.*, 2006). Information about parasitic infection in Southeast Asia and Cambodia is limited due to the lack of quality instruments and methods needed for detections (Dorny *et al.*, 2011; Jeon *et al.*, 2011; Khieu *et al.*, 2014a; Khieu *et al.*, 2014b; Miyamoto *et al.*, 2014). There are few reports that investigated the risk factors of *G. duodenalis* assemblages and the association between *Giardia* assemblages and clinical symptoms of giardiasis.

Table 1.6 Summary of *Cryptosporidium* spp. infection in Asia

Country	Area of Study (Province)	Subjects	Prevalence (%) reported	Methods	References
Nepal	-	children (5 years of age) with acute diarrhea	5.6% (9/160)	-	Shariff <i>et al.</i> (2002)
India	-	1,338 human stool samples with diarrhoea and without diarrhoea	- 3% (40/1,338) - 4.6% in humans who had diarrhoea - 1.2% in humans who had no diarrhea.	microscopic method	Das <i>et al.</i> (2006)
India	-	40 positive samples	- 93.5% (35/40) with <i>C. hominis</i> - 10% (4/40) with <i>C. parvum</i>	PCR-RFLP of the 18S rRNA gene and sequencing were used for genotyping	Das <i>et al.</i> (2006)

			- 2.5 % (1/40) with <i>C. felis</i>		
India	Southern India	53 positive samples	- 81% (43/53) was <i>C. hominis</i> - 12 % (6/53) was <i>C. parvum</i> - 7% (4/53) were <i>C. felis</i> and <i>C. parvum</i> (mouse genotype)	PCR-RFLP of the 18S rRNA gene	Ajjampur <i>et al.</i> (2007)
India	Kolkata	50 child patients (<i>Cryptosporidium</i> positive samples)	<i>C. hominis</i> (49/50) <i>C. meleagridis</i> (2/50) <i>C. felis</i> (1/50)	Multilocus Sequence Typing (MLST)	Gatei <i>et al.</i> (2007)
India	-	48 humans who were infected with HIV	<i>C. hominis</i> , <i>C. parvum</i> , <i>C. felis</i> , <i>C. mulis</i> , <i>C. meleagridis</i>	Multilocus genotyping	Muthusamy <i>et al.</i> (2006)
China	-	252 stool samples from children with diarrhea in a paediatric clinic (169 samples) and an intestinal clinic (83 samples)	13.5% (34/252)	The nested PCR	Liu <i>et al.</i> (2014)

1.7 Development of metabolomics method for the analysis of human faecal samples using *Cryptosporidium*-infected samples.

Metabolomics, the study of the metabolome involves investigation of endogenous and exogenous metabolites. Metabolites play an important role in anabolic and catabolic pathways and as physiological or pathological conditions change so will the metabolites, thus providing potentially important information about the health of an individual organism. Metabolites have been profiled from tissue and bodily fluids, with urine and blood being the most common (Bezabeh *et al.*, 2009). The profiling of faecal metabolites has been shown to be useful in informing the investigation of diseases of the gastrointestinal tract, particularly between the many microbial communities in the tract (Ponnusamy *et al.*, 2011)

Profiling of faecal metabolites is still in its early stages, but should allow for investigation of the complex relationships between gastrointestinal infections/diseases and host health. Gas chromatography – mass spectrometry (GCMS) is a sensitive and reliable method for metabolite analysis (Villas-Boas *et al.*, 2005). GCMS is a very useful tool therefore for the study of faecal samples and their metabolites. Use of an optimised method of extraction is vital to make sure that as many metabolites as possible are harvested from a sample. It is also important to ensure that sample handling and biochemical changes are metabolites from minimised as much as possible. Metabolic extractions from faecal samples have been described in three separate studies. (Ponnusamy *et al.*, 2011; Gao *et al.*, 2009; 2010). Extracting from human faecal water as performed by (Gao *et al.*, 2009; 2010) offered for good metabolite extraction however large amounts of faecal matter were required. Ponnusamy *et al.* (2011) described extraction of metabolites from faecal samples from patients with irritable bowel syndrome (IBS) with little sample manipulation. The GCMS results showed differences between patients with IBS and those without. However, no data were presented on how the method was/could be improved to the optimum standard. Also the study did not take into account the faecal consistencies of subjects, particularly those who were suffering from gastrointestinal infections. Faecal gas headspace extraction and GCMS were used to extract and identify Volatile Organic Compounds (VOCs). Analytical tools including Metab, R (Aggio *et al.*, 2011) and Metaboanalyst (Xia *et al.*, 2009) were subsequently used. Stool samples were obtained from patients with diagnosed/confirmed isolated pathogen infections, together with stool samples from subjects without pathogenic infections which were used as controls. Any significant differences between the VOCs were then analysed.

The study of faecal metabolic profiling is still in its early stages and most studies have used nuclear magnetic resonance (NMR) spectroscopy (Bezabeh *et al.*, 2009; Ponnusamy *et al.*, 2011). Several methods are available for the laboratory diagnosis of *Cryptosporidium* spp. Having unique metabolic pathways due to its lack of mitochondria, it makes an ideal target for the profiling of VOCs. In a previous study about metabolomics, Ng *et al.* (2012) describe a faecal metabolite extraction method for untargeted gas chromatography-mass spectrometry (GCMS) analysis using human faecal samples that include both samples positive for *Cryptosporidium* and negative. The method of extraction takes into account the differing quantities and consistencies required for

clinical diagnosis. Different extraction solvents were used on three different faecal quantities to determine the minimum sample required. The method was validated by untargeted GCMS analysis on eight samples positive for *Cryptosporidium* and eight samples that were negative. This method showed good extraction reproducibility with a relative standard deviation of 9.1%. Multivariate analysis of the GCMS generated dataset showed noticeable profile differences between *Cryptosporidium* positive samples and those which were *Cryptosporidium* negative (Ng *et al.*, 2012).

1.8 AIMS AND OBJECTIVES

Aims

In Africa, the number of molecular epidemiological studies of giardiasis and cryptosporidiosis in humans is small (Cranendonk *et al.*, 2003; Gatei *et al.*, 2003; Morse *et al.*, 2007; Peng *et al.*, 2003). However, there was no evidence of published papers for giardiasis in Malawi as a result there was no information about the incidence and genotyping of *G. duodenalis* in this country. As for cryptosporidiosis research studies in Malawi, there are few reports that have been undertaken on the incidence of infections of *Cryptosporidium* spp. and other protozoa parasites (Cranendonk *et al.*, 2003; Pavone *et al.*, 1990). Moreover, most prevalent information is based on microscopy. Therefore, information on *G. duodenalis* and *Cryptosporidium* species or genotypes is rare. For this reason, we decided to investigate the incidences of *G. duodenalis* and *Cryptosporidium* spp. in Malawi.

The purpose of our study is to better understand the zoonotic transmission potential of Giardiasis and Cryptosporidiosis in human cases from Malawi and Cambodia by using a molecular epidemiological approach and appropriate markers. Our study also assessed the extent of variation in sequences from the isolates within each genotype.

Objectives:

1. To determine the range of *G. duodenalis* genotypes (assemblage prevalence and zoonotic genotypes) and *Cryptosporidium* species infecting children under 5 years of age from diverse geographical regions of Malawi.

2. To investigate the distribution of *G. duodenalis* genotypes (assemblage prevalence and zoonotic genotypes) from faecal human samples in children under 16 years of age with acute diarrhea in different locations in Cambodia.
3. To detect genetic variation among each of the *G. duodenalis* assemblages and compare genetic evolutions of these assemblages isolated from Malawi and Cambodia.
4. To investigate the use of Gas Chromatography and Mass Spectrometry (GCMS) as a novel method for studying the volatile organic compounds in the faeces of patients with *Cryptosporidium* infection in Malawi.

CHAPTER TWO

GIARDIASIS AND CRYPTOSPORIDIOSIS IN MALAWI

2.1 INTRODUCTION

2.1.1 Diarrhoeal diseases in Africa

Diarrhoeal disease is a leading global cause of morbidity and mortality, especially in the developing countries of Asia, Africa and Latin America (Liu *et al.*, 2016). Diarrhoea is a major cause of mortality amongst children. In 2015 it was responsible for 9 percent of all child deaths of those aged five or less. Even though effective treatment is available this equates to about 526,000 child deaths every year. In Malawi, diarrhoea is the sixth leading cause of death in children aged < 5 years (GBD 2015 Mortality and Causes of Death Collaborators, 2015). Many pathogens can cause diarrhoea including bacteria (principally *Campylobacter*, *Salmonella*, *Shigella*, and *Escherichia coli*), viruses (rotavirus, enteric adenovirus, norovirus, sapovirus, astrovirus, and protozoa (*Giardia duodenalis*, *Cryptosporidium* spp. and *Entamoeba histolytica*).

2.1.2 *Giardia duodenalis* and *Cryptosporidium* spp. infection in Africa

Overall, the prevalence of *G. duodenalis* and *Cryptosporidium* spp. in Africa varies in different regions. However, it depends on the methods used for detection, sample type (faeces, duodenal aspirates), host/patient demographics (age groups, gender, immune status (HIV/AIDS patients, cancer patients, transplant patients), sources of samples (urban/rural) and the level of sanitation and hygiene in each community.

In children <5 years of age with diarrhoea, the prevalence of Giardiasis ranged from 10%-29 % by using faecal microscopy in African countries including Ethiopia (Adamu *et al.*, 2013a), Rwanda (Ignatius *et al.*, 2012), Zambia (Siwila *et al.*, 2011), Ghana (Boaitey *et al.*, 2012) and South Ethiopia (Wegayehu *et al.*, 2013a). In Ghana, a total of 19% of HIV seropositive patients with diarrhoea among the various age groups were infected with *G. duodenalis* (Boaitey *et al.*, 2012). However, in high-endemicity areas in East Africa such as Southern highland Rwanda about 60% were found positive for *G. duodenalis* by PCR whereas 19.8% positivity was found using microscopy in

children under five years of age (with diarrhoea and other symptoms) from the Centre Hospitalier Universitaire de Butare (CHUB), several health centers, and Kabutare district hospital (Ignatius *et al.*, 2012). This study showed that this molecular method was of higher sensitivity than conventional methods. It was suggested that the prevalence of *G. duodenalis* in these areas was probably highly underestimated due to the use of the low sensitive microscopic method (Ignatius *et al.*, 2012). Using multi-locus genotyping analysis (SSU-rDNA *bg*, *gdh* and *tpi*), the *G. duodenalis* assemblage B was the most predominant in Africa (Ankarklev *et al.*, 2012; Di Cristanziano *et al.*, 2014; Ignatius *et al.*, 2012). Gender was not associated with *G. duodenalis* infections and there was no significant difference of infection among the age groups (Wegayehu *et al.*, 2013b). There are limited amount of *Giardia* data for other African countries. In particular, there is currently no literature describing *Giardia* in Malawi.

The prevalence of Cryptosporidiosis in Africa ranged from 12-30.5% in children with diarrhoea <5 years of age by using microscopy and the modified Ziehl-Neelsen (MZN) staining method. The MZN stain also known as the hot method of Acid Fast Bacteria (Bacilli) or AFB staining. This method is used to identify those bacteria which are acid-fast, mainly *Mycobacteria*. These acid fast organisms (fast = holding capacity) are those which are able to keep the primary stain when treated with an acid. It is also useful to confirm the presence of oocysts of coccidian parasites such as *Cryptosporidium* spp., *Isospora belli* and *Cyclospora cayetanensis*. Oocysts of these coccidian parasites are also acid-fast (Moyo *et al.*, 2011; Samra *et al.*, 2012; Mbae *et al.*, 2013). However, the prevalence of Cryptosporidiosis is higher in immunosuppressed infants; 60.2% were infected with *C. parvum* (Abdel-Hafeez *et al.*, 2012). In adult patients with HIV/AIDS and diarrhoea, oocysts of *Cryptosporidium* spp. were detected in 8.5% in Ethiopia (Adamu *et al.*, 2010) and 11% of adult patients with diarrhoea admitted to the Medical Wards in Queen Elizabeth Central Hospital (QECH) in Blantyre, Malawi had *C. parvum* (Cranendonk *et al.*, 2003). This demonstrated that adult patients with diarrhoea had lower prevalence than in children <5 years of age with diarrhoea and also in immunosuppressed infants. In HIV-infected adult patients, the prevalence of Cryptosporidiosis ranged from 13.3 %-19 % in West Cameroon (Nkenfou *et al.*, 2013) and in Ethiopia (Zeynudin *et al.*, 2013) 2.5% were infected with *C. parvum* in the non-HIV infected patients.

C. parvum was the most predominant (19%) in HIV/AIDS patients from West Cameroon (Nkenfou *et al.*, 2013) compared to other parasites.

Using molecular techniques, the prevalence of *Cryptosporidium* spp. in humans from Western Uganda was 32.4% by a nested PCR, of the *Cryptosporidium* oocyst wall protein (COWP) gene (Salyer *et al.*, 2012). Immunosuppressed infants had the highest prevalence of *Cryptosporidium* spp. in Egypt (60.2%)(Abdel-Hafeez *et al.*, 2012), followed by children <5 years of age with diarrhoea (12-30.5 %), 12.2% in South Africa (Samra *et al.*, 2012), 18.9% in Tanzania (Moyo *et al.*, 2011), 30.5% in Nairobi, Kenya (Mbae *et al.*, 2013). For HIV patients (13.3 %-19 %), in West Cameroon (Nkenfou *et al.*, 2013) and in Ethiopia (Zeynudin *et al.*, 2013). Adult patients with diarrhoea (8.5-11%), 8.5% in Ethiopia (Adamu *et al.*, 2010) and 11% of adult patients in Blantyre, Malawi (Cranendonk *et al.*, 2003), respectively. For genotyping of *Cryptosporidium* spp., the most predominant species was *C. hominis* (95.3%, 41/43), followed by *C. parvum* (4.6%, 2/43) (Morse *et al.*, 2007). For other species, *C. meleagridis*, *C. andersoni* and also mixed infection (*C. hominis*/*C. parvum*) were also found in humans (Morse *et al.*, 2007; Peng *et al.*, 2003).

It was concluded that *G. duodenalis* and *Cryptosporidium* spp. were the most common protozoal causes of diarrhoeal disease in African countries (Boaitey *et al.*, 2012; McElligott *et al.*, 2013). *Cryptosporidium* spp. was most common followed by *G. duodenalis*. Many researchers have been more interested in studying Cryptosporidiosis than Giardiasis which may be due to *Cryptosporidium* spp. being significantly more often associated with diarrhoea compared to giardiasis (Boaitey *et al.*, 2012; McElligott *et al.*, 2013; Mbae *et al.*, 2013). Gender was not associated with *G. duodenalis* infections and there was no significant difference of infection among the age groups (both age and sex were not associated with *G. duodenalis* infections) (Adamu *et al.*, 2013b).

In Africa, the number of molecular epidemiological studies of giardiasis and cryptosporidiosis in humans is small (Gatei *et al.*, 2003; Peng *et al.*, 2003; Cranendonk *et al.*, 2003; Morse *et al.*, 2007). While a study of *Cryptosporidium* spp. from Malawi was previously published (Gatei *et al.*, 2003; Peng *et al.*, 2003; Cranendonk *et al.*, 2003; Morse *et al.*, 2007). There were no published papers for

giardiasis in Malawi. Moreover, most information is based on microscopy. The current study focused on diarrhoea-causing parasitic protozoa, *G. duodenalis* and *Cryptosporidium* spp. only.

2.2 AIM OF STUDY

The aim of this study is to determine the range of *Giardia duodenalis* genotypes (assemblage prevalence) and *Cryptosporidium* species identified in children with diarrhoea in Malawi. The samples analysed were recovered from children in Malawi under 5 years of age, living in different geographical regions as part of a previous study of rotavirus infections. Three regions which were selected comprise Karonga (with a small rural hospital), Lilongwe (a large urban hospital) and Mangochi (a medium sized rural hospital). The three regions were selected so that they were representative of the northern, central and southern parts of Malawi respectively. Moreover, it was the best way to study genetic diversity as the three locations show differences in the size of the hospital, geographic location and development (both rural and urban areas were included). The hypotheses of our study are that: 1) there will be anthroponotic species in urban areas 2) there will be zoonotic species in rural areas 3) there will be mixed infection (anthroponotic and zoonotic species) in urban and rural areas.

Appropriate markers and/or multilocus genotyping was used to identify assemblages and sub-assemblages of *G. duodenalis*. Our study also assessed the extent of variation in sequences from the isolates within each genotype. Faecal samples were screened for both parasites (*Giardia* and *Cryptosporidium*) microscopically using the Direct Fluorescence Assay (DFA) and parasite isolates were typed by using multilocus genotyping, Nested-PCR and/or sequencing of SSU-rRNA, Triose phosphate isomerase gene (*tpi* gene), β -giardin (*bg* gene), and glutamate dehydrogenase genes (*gdh* gene) for *Giardia*. For genotyping of *Cryptosporidium* spp., Nested-PCR and/or sequencing of SSU-rRNA was used. DNA sequence analysis of the 60 kDa glycoprotein (GP60) gene was used for subtyping of *C. parvum* and *C. hominis*.

1. The questions to be addressed in this Chapter include: How does the prevalence of *Cryptosporidium* and *Giardia* infection vary by age?
2. How does the prevalence of *Cryptosporidium* spp. and *Giardia* infection vary when using different genes or markers for their detection?

2.3 MATERIALS AND METHODS

2.3.1 Source of samples

The study was carried out in three regions of Malawi (Figure 2.1), namely, Karonga, Lilongwe, and Mangochi to represent northern, central and southern part of Malawi, respectively.

The faecal samples used here were collected as part of previous studies on rotavirus in Malawi (Cunliffe *et al.*, 2009a; Cunliffe *et al.*, 2009b). From July 2005 through June 2007, acute diarrhoeic faecal human samples were collected from children under 5 years of age who presented themselves to the clinic (outpatients) at the Chilumba Rural Hospital in Karonga (a small rural hospital); who presented themselves to the clinic and were admitted to the paediatric wards (outpatients and inpatients) at the Kamuzu Central Hospital in urban Lilongwe (a big urban hospital); and who were admitted to the paediatric wards (inpatients) at the Mangochi District Hospital in rural Mangochi (a medium rural hospital). The clinical information such as gender, date of collection and date of birth (see Appendix 1; Table 1A - Table 1C) was recorded by researchers who collected the faecal specimens from Malawi. Stool samples were collected from each case-patient and were kept in sterile containers at -80°C until rotavirus detection was undertaken. Nucleic acid extracted from the faecal specimens had been used for rotavirus detection and characterisation in a previous study (Cunliffe *et al.*, 2009a; Cunliffe *et al.*, 2009b). Rotavirus was detected previously by ELISA (Cunliffe *et al.*, 2007).

However, they have not been examined for *Giardia* and *Cryptosporidium*, therefore the samples were screened for the presence of these parasites. The samples were defrosted prior to examination several times.

The microtubes contained stool samples which were checked to see if they had enough volume to extract DNA. The samples which had the volume of about 0.5-0.75 ml were chosen to be studied. Moreover, the consistency of the faeces was also checked to make sure that the samples in each tube were not too dry (loose or watery stool samples were used). Samples should be in a good enough condition in order to extract DNA.



Figure 2.1 Map of Malawi showing the regions of Karonga (1), Lilongwe (2) and Mangochi (3).
Taken from: https://commons.wikimedia.org/wiki/File:Malawi_map.png

2.3.2 Direct Fluorescence Assay

The Direct Fluorescent Antibody Test (DFA) is a technique used for detecting specific antigens. Antibodies conjugated with a fluorescent dye is placed in the wells of slides and combines slides with specific antigen (microbes and some protozoa, *Giardia* sp. or *Cryptosporidium* spp.) and allowed to bind, then antigens can be detected using a fluorescent microscope. If the antigen is present it will emit light. It is both a sensitive and specific test. For this study, the direct immunofluorescence test (DFA) for *Giardia* and *Cryptosporidium* spp. was performed using a Merifluor® *Cryptosporidium/Giardia* kit (Meridian Biosciences, UK) according to the manufacturer's instructions. First, a drop of the original faecal sample was taken by using a transfer

loop and then the sample was spread over the entire wells of the slide and left to air dry on a flat surface at room temperature for 30 minutes. After that, one drop of detection reagent and one drop of counterstain were added to each well, the reagents were mixed softly with an applicator stick and spread over the entire well carefully, trying not to scratch the surface of the slide. Slides were placed in a moist chamber and left in the dark at room temperature for 30 minutes. Slides were washed carefully with washing buffer until the reagents and counterstaining disappeared. Care was taken to avoid disturbing the samples and to minimise the risk of cross contamination of specimens (for example during rinsing; the slides were not submerged). Excess buffer was removed by touching the edge of the slide onto absorbent paper and trying to avoid allowing the slide to dry. After that, one drop of mounting media was added to each well and a cover slide with a 22x64 coverslip applied. Finally, slides were analysed with a Zeiss standard 14 microscope fitted with epi-fluorescence IV FI condenser (FITC excitation, 490nm wavelength, light source is HbO 50W bulb) to search for signs of fluorescent bodies by scanning each slide well completely. A Nikon Eclipse 80i, NIS-Elements Basic Research v.3.1 was used for taking photographs.

2.3.3 DNA extraction from faecal samples

DNA was extracted from DFA positive stool samples using the QIAamp DNA Stool mini kit (QIAGEN®) according to the manufacturer's instructions, with minor modifications. Generally, this kit is used for DNA extraction from enteric pathogens in faecal specimens that can be lysed easily (i.e. without oocyst wall or spores). After the preliminary experiments, only a small amount of DNA was obtained from stool extraction, possibly due to the hard and thick wall of *Cryptosporidium* oocysts. Therefore, the procedure was modified to include the addition of five freeze-thaw cycles and putting three glass beads into the tubes before extraction in order to break the oocyst walls of *Cryptosporidium* spp. (Gatei *et al.*, 2003).

The instructions in the kit insert stipulated that the first step is adding lysis buffer onto 200 µl of liquid in a 2 ml micro centrifuge tube, and incubate at 95°C for 10 minutes to extract DNA. However, 1.2 ml of lysis buffer (ASL, buffer provided) was added and used to extract total genomic DNA during the freeze/thaw method. (200 µl of ASL lysis buffer was used to suspend the stool from frozen specimens and the small volume of buffer used in the freeze-thaw to avoid rupture or

overflow during the freeze thaw steps). The mixture was vortexed until thoroughly homogenised. The samples were heated for 10 minutes at 95 °C, glass beads were added and then subjected to five rounds of freeze-thawing (Liquid nitrogen was used for freezing, and thawing was carried out at 95 °C in the heating block). After the final freeze-thaw, an additional 1.2 ml of ASL lysis buffer was added into the faeces mixture. The tubes were vortexed for 1 minute until the sample completely homogenized, then incubated for 10 minutes at 70°C. The tubes were vortexed for 15 seconds then centrifuged at 1300 x g for 1 minute. 1.2 ml of supernatant was transferred into a new 2 ml micro centrifuge tube, a tablet of InhibitEX® was added into the sample tubes to adsorb DNA degrading substances and PCR inhibitors. It was then vortexed until the tablet was totally suspended, then incubated at the room temperature for 3 minutes. Samples were centrifuged at 1300 x g for 3 minutes, the supernatant was transferred into a new 1.5 ml micro tube and centrifuged at 1300 x g for 3 minutes. In order to get rid of protein thoroughly, 200µl of supernatant from the previous step was mixed with 15µl proteinase K and 200µl Buffer AL into another 1.5 ml micro tube. The samples were vortexed for 15 seconds, incubated at 70°C for 10 minutes and centrifuged briefly after incubation. Next, 200µl of absolute ethanol was added, vortexed to mix and centrifuge briefly before applying to a QIAmp spin column and centrifuging at 1300 x g for 1 minute. Next, 500µl of buffer AW1 was added to the column and centrifuged at 1300 x g for 1 minute. After that, the spin column was placed in another collection tube. 500µl of buffer AW2 was added and centrifuged at 1300 x g for 3 minutes (to make sure all of AW2 was removed samples were centrifuged again at 1300 x g for 1 minute). To elute the DNA, 50µl of buffer AE was added directly onto the membrane, incubated at room temperature for 5 minutes and centrifuged at 1300 x g for 1 minute.

2.3.4 Multilocus genotyping of *Giardia* and *Cryptosporidium*

2.3.4.1 Genotyping and subtyping of *Giardia*

Primers were selected from published papers for genotyping and subtyping of *Giardia* (Table 2.1). All primers (except primers for GP60) were ordered from Eurofins MWG Operon (Wesway Estate, Acton, UK). GP60 primers were available from within the Institute of Infection and Global Health.

Table 2.1 Lists of primers for genotyping and subtyping of *Giardia*

Target Gene	Name of Primers & Sequences (5'-3')	Product size (bp)	Reference(s)
SSU	RH11 : CATCCGGTCGATCCTGCC RH4 : AGTCGAACCCTGATTCTCCGCCAGG	292	Hopkins <i>et al.</i> (1997)
<i>tpi</i> (general)	AL3543 : AAATATGCCTGCTCGTCG AL3546 : CAAACCTTITCCGCAAACC	605	Sulaiman <i>et al.</i> (2003)
<i>tpi</i> (assemblage-specific = A)	TPiAf : CGCCGTACACCTGTC A TPiAr : AGCAATGACAACCTCCTTCC	332	Geurden <i>et al.</i> (2008)
<i>tpi</i> (assemblage-specific = B)	TPiBf: GTTGTTGTTGCTCCCTCCTTT TPiBr: CCGGCTCATAGGCAATTACA	400	Geurden <i>et al.</i> (2009a); Geurden <i>et al.</i> (2009b)
<i>bg</i>	G7 : AAGCCCGACGACCTCACCCGCAGTGC G759 : GAGGCCGCCCTGGATCTTCGAGACGAC	753	Cacciò <i>et al.</i> (2002); Lalle <i>et al.</i> (2005)
<i>bg</i>	BGf: GAACGAACGAGATCGAGGTCCG BGr : CTCGACGAGCTTCGTGTT	511	Cacciò <i>et al.</i> , (2002); Lalle <i>et al.</i> (2005)
<i>gdh</i>	GDH1 : TTCCGTRTYCAGTACAACCTC GDH2 : ACCTCGTTCTGRGTGGCGCA	700	Cacciò <i>et al.</i> (2008)
<i>gdh</i>	GDH3: ATGACYGAGCTYCAGAGGCACGT GDH4 : GTGGCGCARGGCATGATGCA	530	Cacciò <i>et al.</i> (2008)

2.3.4.2 PCR amplification of small subunit-rDNA gene (18S SSU-rDNA)

A PCR product that was 292 bp long was amplified by using primers RH11 and RH4 (Hopkins *et al.*, 1997) in standard mixtures of 25 µl containing 10 pmol of each SSU rRNA specific primer, 400 µM of each dNTP, 2.5 µl 10X buffer with 15mM MgCl₂, 5 µl Q Solution, 1.25 U *Taq* DNA polymerase and 2 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 20 seconds, 59°C for 20 seconds, and 72°C for 30 seconds, were performed; an initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included. The use of hot start PCR enables more specific and greater yields of the DNA, by using specific antibodies at lower temperatures and blocking the *Taq* polymerase. All PCR reagents were from QIAGEN®

(Manchester, UK) and dNTPs from Sigma-Aldrich® (Manchester, UK). The PCR reactions were carried out in a DNA Engine Dyad® Peltier Thermal Cycler (MJ Research Inc).

2.3.4.3 Triose phosphate isomerase (*tpi*) gene

The *tpi* gene was amplified in an assemblage specific nested reaction. A primary reaction of PCR product that was 605 bp long was amplified by using primers AL3543 and AL3546 (Sulaiman *et al.*, 2003) in standard mixtures of 25 µl containing 10 pmol of each primer, 400 µM of each dNTP, 2.5 mM MgCl₂, 2.5 µl 10X buffer, and 1.25 U of *Taq* DNA polymerase and 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 45 seconds, 58°C for 45 seconds, and 72°C for 1 minute, were performed; an initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included.

2.3.4.3.1 *Giardia* Assemblage A specific PCR

Using the 605 bp fragment as the template DNA, a 332 bp PCR product was amplified by using Assemblage A specific primers TPiAf and TPiAr (Geurden *et al.*, 2008) in standard mixtures of 25 µl containing 10 pmol of each specific primer, 400 µM of each dNTP, 2.5 mM MgCl₂, 10X buffer, 1.25 U of *Taq* DNA polymerase, and 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 45 seconds, 58°C for 45 seconds, and 72°C for 45 seconds, were performed; an initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included.

2.3.4.3.2 *Giardia* Assemblage B specific PCR

Using the 605 bp fragment as the template DNA, a 400 bp PCR product of was amplified by using Assemblage B specific primers TPiBf and TPiBr (Geurden *et al.*, 2009a; Geurden *et al.*, 2009b) in standard mixtures of 25 µl containing 10 pmol of each specific primer, 400 µM dNTP, 2.5 mM MgCl₂, 10X buffer, 1.25 U of *Taq* DNA polymerase and 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 45 seconds, 62°C for 45 seconds, and 72°C for 45 seconds, were performed; an initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included.

2.3.4.4 *β-giardin (bg)* gene

The gene was amplified as a nested reaction; a primary reaction to generate a PCR product of 753 bp was amplified by using primers G7 and G759 (Caccio *et al.*, 2002; Lalle *et al.*, 2005) in standard mixtures of 25 µl containing 400 nmoles of each specific primer, 200 µM dNTP, 2.5 mM MgCl₂, 2.5 µl 10X buffer and 1.25 U of *Taq* DNA polymerase and 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 30 seconds, 65°C for 30 seconds, and 72°C for 1 minute, were performed; an initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included. For the secondary PCR step, a 511 bp PCR product was amplified by using 2 µl of the primary PCR product and primers BGf and BGr (Caccio *et al.*, 2002; Lalle *et al.*, 2005) in standard mixtures of 25 µl containing 10 pmol of each specific primer, 400 µM dNTP, 10X buffer, 1.25 U of *Taq* DNA polymerase and 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 30 seconds, 64°C for 30 seconds and 72°C for 1 minute were performed. An initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes was included.

2.3.4.5 Glutamate dehydrogenase (*gdh*) gene

The GDH gene was amplified as a nested reaction; a primary reaction of PCR product that was 700 bp long was amplified by using primers GDH1 and GDH2 (Caccio *et al.*, 2008) in standard mixtures of 25 µl containing 10 pmol of each specific primer, 400 µM dNTP, 2.5 mM MgCl₂, 2.5 µl 10X buffer, 1.25 U of *Taq* DNA polymerase and 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 1 minute were performed. An initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included and held at 4 °C. For the secondary PCR step, a 530 bp PCR product was amplified by using 2 µl of the primary PCR product and primers GDH3 and GDH4 (Caccio *et al.*, 2008) in standard mixtures of 25 µl containing 10 pmol of each specific primer, 400 µM dNTP, 2.5 mM MgCl₂, 10X buffer, 1.25 U of *Taq* DNA polymerase and 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 30 seconds, 68°C for 30 seconds and 72°C for 1 minute, were performed. An initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included and hold at 4 °C.

In order to ensure that the amplification was successful and to confirm the size of the PCR products, electrophoresis was used. Briefly, a 1.5% agarose gel was prepared in a horizontal plastic chamber, weighed DNA grade-agarose was heated to dissolve in Tris-Borate-EDTA (TBE, 45 mM Tris-borate and 1 mM EDTA) buffer and cooled to 50 °C before pouring into the tape-sealed gel tray. Appropriate gel comb was inserted into the gel and the gel was allowed to set at 26 °C for 45 minutes. The 1.5% agarose gel stained with SYBR Safe DNA gel stain (Invitrogen™) was loaded with 5 µl of amplified product mixed with 2 µl of loading dye and 5 µl of 100bp DNA ladder (Thermo Fisher Scientific). Electrophoresis was carried out at 110 volts in the 1XTBE buffer for 45 minutes or until the dye front reached 75% across the gel. The gel was visualized under a UV transilluminator and then the photographs were taken with Image Master (Bio-Rad Laboratories Ltd, UK).

2.3.5 Genotyping and subtyping of *Cryptosporidium* spp.

2.3.5.1 Primers for genotyping and subtyping of *Cryptosporidium* spp.

Primers were selected from published papers for genotyping and subtyping *Cryptosporidium* spp. (Table 2.2).

Table 2.2 Lists of primers for genotyping and subtyping of *Cryptosporidium* spp.

Target Gene	Name of Primers & Sequences (5'-3')	Expected product size (bp)	Reference (s)
SSU-rRNA	CPSSU1 : TTC TAG AGC TAA TAC ATG CG CPSSU2 : CCC ATT TCC TTC GAA ACA GGA	1325	Xiao <i>et al.</i> (2001b)
SSU-rRNA	CPSSU3 : GGA AGG GTT GTA TTT ATT AGA TAA AG CPSSU4 : AAG GAG TAA GGA ACA ACC TCC A	826-864	Xiao <i>et al.</i> (2001b)
GP60	GP601 : ATA GTC TCC GCT GTA TTC GP602 : GGA AGG AAC GAT GTA TCT	800-850	Alves <i>et al.</i> (2003)
GP60	GP603 : TCC GCT GTA TTC TCA GCC GP604 : GCA GAG GAA CCA GCA TC	400-450	Alves <i>et al.</i> (2003)

2.3.5.2 SSU-rRNA gene

SSU rRNA nested-PCR was performed. Briefly, for the primary PCR step, a PCR product that was 1,325 bp long was amplified by using primers CPSSU1 and CPSSU2 (Xiao *et al.*, 2001b) in standard mixtures of 25 µl containing 10 pmol of each SSU rRNA specific primer, 400 µM dNTP, 9.5 mM MgCl₂, 2.5 µl 10X buffer and 1.5 U *Taq* DNA polymerase, 1% BSA, 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 45 seconds, 58°C for 45 seconds and 72°C for 1 minute, were performed. An initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included and held at 4 °C. For the secondary PCR step, a PCR product that was 826 bp long was amplified by using 1 µl of the primary PCR product and primers CPSSU3 and CPSSU4 (Xiao *et al.*, 2001b). A total of 35 cycles, each consisting of 94°C for 45 seconds, 58°C for 45 seconds and 72°C for 1 minute were performed. An initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included and held at 4 °C.

2.3.5.3 Restriction Fragment Length Polymorphism (RFLP) of the SSU PCR

For species diagnosis, *SspI* (Sigma-Aldrich) was used to digest the SSU-PCR product (Xiao *et al.*, 2004). 10 µl of the secondary PCR product of SSU-rRNA gene was digested in a 25 µl reaction mixture consisting 20 U of *SspI* (20 U/reaction), 2.5 µl of 10X buffer. The reactions were incubated at 37 °C for 1 hour and visualised on an agarose gel.

For genotyping of *C. parvum*, *VspI* (Promega) was used (Xiao *et al.*, 2004). 10 µl of the secondary PCR product of SSU-rRNA gene was digested in a 25 µl reaction mixture consisting, 20 U of *VspI* (20 U/reaction), 2.5 µl of 10X buffer. The reactions were incubated at 37 °C for 1 hour and visualised on an agarose gel. Species and genotypes were determined by banding patterns and confirmed by DNA sequencing of the PCR products (Table 2.3).

Table 2.3 Patterns of RFLP digestion for *Cryptosporidium*

Species	Genotype	<i>Ssp</i> I fragment size (bp)	<i>Vsp</i> I fragment size (bp)
<i>C. parvum</i>	‘human’ genotype (<i>C. hominis</i>) type 1	449, 254, 111	526, 104, 102, 70
<i>C. parvum</i>	‘bovine’ genotype (<i>C. parvum</i>) type 2	449, 254, 108	628, 104
<i>C. parvum</i>	Dog	417, 254, 105	633, 102
<i>C. parvum</i>	Monkey	461, 254, 109, 11	559, 104
<i>C. meleagridis</i>		449, 254, 108	456, 171, 104
<i>C. muris</i>		448, 385	731, 102
<i>C. felis</i>		426, 390, 33, 15	476, 182, 104

2.3.5.4 GP60

A fragment of the GP60 gene (800-850 bp) was amplified with the primers GP601 and GP602 (Alves *et al.*, 2003) in standard mixtures of 25 µl containing 10 pmol of each primer, 400 µM dNTP, 4.5 mM MgCl₂, 2.5 µl 10X buffer and 2.5 U *Taq* DNA polymerase, 1% BSA, 1 µl of DNA template. A total of 35 cycles, each consisting of 94°C for 45 seconds, 50°C for 45 seconds, and 72°C for 1 minute, were performed. An initial hot start at 94°C for 3 minutes and a final extension step at 72°C for 10 minutes were also included. For the secondary PCR step, a PCR product that was 400-450 bp long was amplified by using 1 µl of the primary PCR product and primers GP603 and GP604 (Alves *et al.*, 2003). The GP60 gene is hypervariable and has a number of repeat sequences so the product sizes for the different species are different; this property is used to type the *Cryptosporidium* spp. The PCR cycling conditions were identical to the conditions used for the primary PCR step.

2.3.6 Purification of PCR products

All PCR products were purified for sequencing using a QIAquick® PCR Purification kit (QIAGEN®), according to the manufacturer's instructions. Briefly, 5 volumes of buffer PB were added to 1 volume of PCR product and mixed carefully. The sample was added to the QIAquick column and centrifuged at 1300 x g for 1 minute. Then the flow-through was discarded and placed in the column in the same collection tube. After that, 0.75 ml buffer PE was added to the QIAquick column and for 1 minute. Then, the flow-through was discarded, the column was placed in a new collection tube and centrifuged at 1300 x g for a further 1 minute. Lastly, the column was placed in a 1.5 ml micro centrifuge tube, a further 50µl was added for increased DNA concentration, and then the tube was left at the room temperature for 1 minute of PCR water to the centre of the membrane and centrifuged at 1300 x g for 1 minute.

2.3.7 Sequencing of PCR products and DNA Sequence analysis

The purified PCR products were sequenced in both directions (forward and reverse) using an Applied Biosystems® 3730 DNA Analyser at the Core Genomics Facility, University of Sheffield, UK. Sequenced products were compared to homologous sequences in the GenBank database using Basic Local Alignment Search Tool (BLAST) (Zhang *et al.*, 2000) to determine the parasites assemblage/species. BioEdit ver. 7.2.0 (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) was the computer software used to edit and assemble the chromatograms of the DNA sequences (contigs).

2.3.8 Statistical analyses

Associations between *Giardia* and *Cryptosporidium* infection and children's gender, age and rotavirus co-infection were analysed. In order to analyse the correlation among those as mentioned above, statistical analyses were performed using SPSS (version 7.0). The age data were divided into three groups (<1 year, 1-2 years, and 2-4 years). For *Rotavirus*, co-infection were categorised into two groups for the presence of or an absence of *Rotavirus* infection. Comparison of proportions of assemblage A and B were tested by Fisher's Exact test. The level of statistical significance was set at $p < 0.05$.

2.3.9 Ethical approval for the studies

The primary rotavirus study (including sample collection and testing for diarrhoeal pathogens) was approved by College of Medicine Research Ethics Committee, University of Malawi.

2.4 RESULTS

2.4.1 Direct Fluorescence Assay for *G. duodenalis* and *Cryptosporidium* spp.

One hundred and thirty two faecal samples from Malawi were stained by DFA and examined under oil immersion using fluorescence microscopy. The overall prevalence of *G. duodenalis*, *Cryptosporidium* spp. and mixed infection of both species by DFA were 11.4% (15/132), 23.5% (31/132) and 3% (4/132), respectively. The prevalence of *G. duodenalis* was significantly lower than the prevalence of *Cryptosporidium* spp. (Fisher's exact test, $p = 0.014$). There were significant differences between the prevalence of *G. duodenalis* and mixed infection (Fisher's exact test, $p = 0.008$) and the prevalence of *Cryptosporidium* spp. and mixed infection (Fisher's exact test, $p = 0.001$).

The prevalence of *G. duodenalis* in samples collected from Karonga, Lilongwe, and Mangochi was 14.3% (5/35), 6.9% (5/72) and 20% (5/25) respectively. The prevalence of *Cryptosporidium* spp. in samples collected from Karonga, Lilongwe, and Mangochi were 17.1% (6/35), 23.6% (17/72) and 32% (8/25) respectively. Moreover, the prevalence of mixed infections found in Lilongwe and Mangochi were 4.2% (3/72), and 4% (1/25), respectively. No mixed infections were identified in Karonga. There were no significant differences between the prevalence of *G. duodenalis* and *Cryptosporidium* spp. between each region in Malawi (Fisher's exact test, $p > 0.05$). DFA and fluorescence microscopy identified positive samples as having *G. duodenalis* or *G. duodenalis* like-cysts and *Cryptosporidium* spp. or *Cryptosporidium* spp. like-oocysts. (Figure 2.2 & Table 2.4).

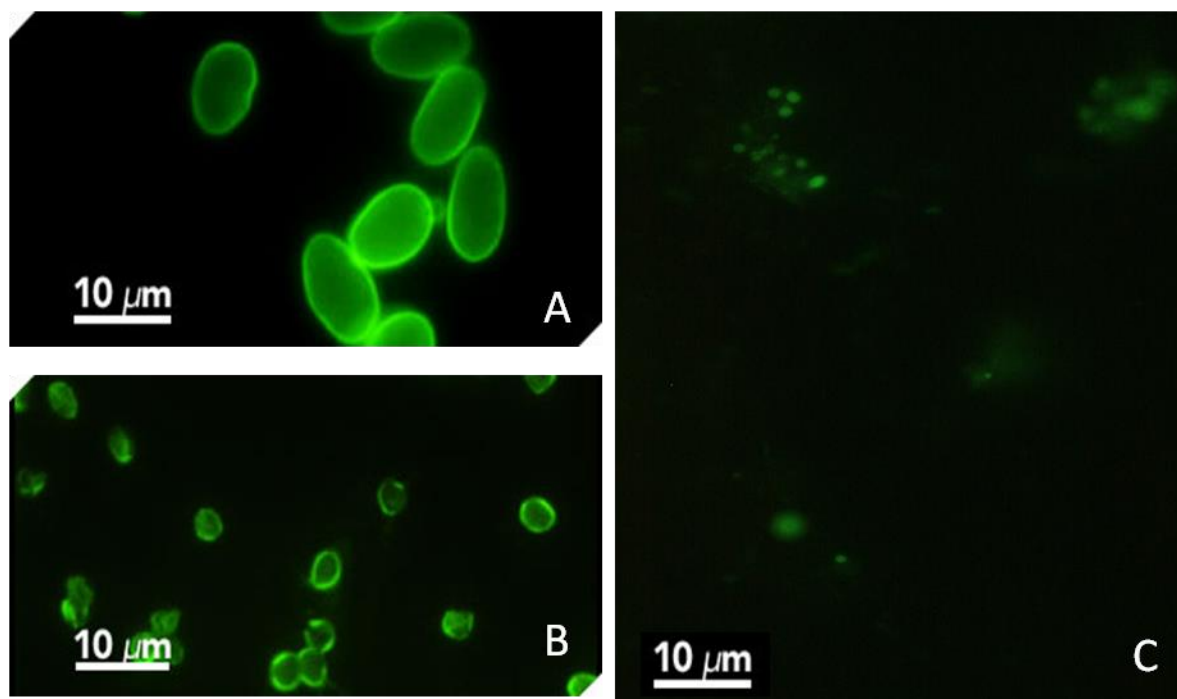


Figure 2.2 *G. duodenalis* cysts by DFA: The *G. duodenalis* cysts are oval in shape, with a bright apple green colour, approximately 8-12 µm long (Panel A). *Cryptosporidium* oocysts are round to slightly oval in shape with a bright apple green colour, approximately 2-6 µm in diameter (Panel B) negative result (neither *G. duodenalis* nor *Cryptosporidium* spp. (Panel C).

Table 2.4 The prevalence of *Giardia* and *Cryptosporidium* infections in children with diarrhoea under 5 years of age from three regions in Malawi

Regions	DFA		
	<i>G. duodenalis</i>	<i>Cryptosporidium</i> spp.	Mixed infection (<i>Giardia</i> + <i>Crypto</i>)
Karonga	14.3% (5/35)	17.1% (6/35)	0% (0/35)
Lilongwe	6.9% (5/72)	23.6% (17/72)	4.2% (3/72)
Mangochi	20% (5/25)	32% (8/25)	4% (1/25)
Total prevalence	11.4% (15/132)	23.5% (31/132)	3% (4/132)

2.4.2 Multilocus genotyping and subtyping of *G. duodenalis*

2.4.2.1 Analysis of the small subunit-rDNA gene (18S SSU -rRNA)

The 292 bp fragment of the SSU-rRNA locus of 200 DNA human samples were amplified. The overall prevalence of *G. duodenalis* in Malawi by 18S SSU-rRNA was 28% (56/200). The prevalence of *G. duodenalis* by 18S SSU-rRNA was significantly different to the prevalence of *G. duodenalis* by *bg* (Fisher's exact test, $p = 0.001$). However, the prevalence of *G. duodenalis* by 18S SSU-rRNA was not significantly different to the prevalence of *G. duodenalis* by *tpi* and *gdh* (Fisher's exact test, $p > 0.05$). There were no significant differences between the prevalence of *G. duodenalis* by 18S SSU-rRNA among those by *tpi* and *gdh* (Fisher's exact test, $p > 0.05$).

The prevalence of *G. duodenalis* recorded from patient samples collected in Karonga, Lilongwe, and Mangochi were 23.4%, 30.6% and 29.7% respectively (Table 2.5). There were no significant differences between the prevalence of *G. duodenalis* by 18S SSU-rRNA among those from each region in Malawi (Fisher's exact test, $p > 0.05$).

Table 2.5 The detection of *G. duodenalis* in children under 5 year of age from three regions in Malawi using SSU-rRNA PCR

Source	SSU-rRNA	<i>tpi</i>	<i>Bg</i>	<i>gdh</i>
Karonga	23.4% (15/64)	23.4% (15/64)	17.2% (11/64)	21.9% (14/64)
Lilongwe	30.6 % (22/72)	30.6 % (22/72)	9.7% (7/72)	20.83% (15/72)
Mangochi	29.7 % (19/64)	29.7 % (19/64)	3.1% (2/64)	25.5% (13/64)
Total	28 % (56/200)	28% (56/200)	10% (20/200)	21% (42/200)

2.4.2.2 Analysis of the Triose Phosphate Isomerase gene (*tpi* gene)

The 332 bp fragment of the *tpi* gene (Assemblage A) and the 400 bp fragment of the *tpi* gene (Assemblage B) of 200 samples were amplified. The prevalence of *G. duodenalis* recorded from samples collected in Karonga, Lilongwe, and Mangochi were 23.4%, 30.6% and 29.7% respectively. The overall prevalence of *G. duodenalis* by the *tpi* gene was 28% (Table 2.5). The prevalence of

G. duodenalis by *tpi* was significantly different to the prevalence of *G. duodenalis* by *bg* (Fisher 's exact test, $p = 0.00$). However, the prevalence of *G. duodenalis* by *tpi* was not significantly different to the prevalence of *G. duodenalis* by 18S SSU-rRNA and *gdh* (Fisher 's exact test, $p > 0.05$). There were no significant differences between the prevalence of *G. duodenalis* by 18S SSU-rRNA among those by *tpi* and *gdh* (Fisher 's exact test, $p > 0.05$). There were no significant differences between the prevalence of *G. duodenalis* i by *tpi* among those from each region in Malawi (Fisher 's exact test, $p > 0.05$).

Out of 56 *Giardia* positive samples from three regions in Malawi, 42 (75%) were successfully sequenced and genotyped. The results of the sequencing analysis matched the GenBank reference (Table 2.6). In the genotyping of *G. duodenalis*, the infection rates of assemblage B and mixed infection of assemblage A and B were significantly higher than assemblage A. The prevalence of Assemblage A by *tpi* was significantly different to the prevalence of Assemblage AB by *tpi* (Fisher 's exact test, $p = 0.013$). However, there were no significant differences between the prevalence of Assemblage A, Assemblage B and *G. duodenalis* determined by *tpi* (Fisher 's exact test, $p > 0.05$) (Table 2.6). By using the *tpi* gene, 100% (19/19) belonged to the sub-assemblage AII from all assemblage A. Among Assemblage B parasites, 87% (26/30) belonged to sub-assemblages BIII, 7% (2/30) displayed sub-assemblage BIV, 2% (1/30) belonged to sub-assemblage B (VB906855) and 2% (1/30) showed sub-assemblage other B.

Table 2.6 *G. duodenalis* assemblages (*tpi* gene) in children under 5 year of age from three regions in Malawi

Source	Assemblage A	Assemblage B	Assemblage AB	Seq. not determined	Total
Karonga	0% (0/15)	40 % (6/15)	20% (3/15)	40% (6/15)	15
Lilongwe	13.6 % (3/22)	27.3% (6/22)	45.5% (10/22)	13.6% (3/22)	22
Mangochi	26.3%(5/19)	15.8% (3/19)	31.6% (6/19)	26.3% (5/19)	19
Total	14.3 % (8/56)	26.8% (15/56)	33.9% (19/56)	25% (14/56)	56

2.4.2.3 Analysis of the β -giardin gene

The 511 bp fragment of the β -giardin (*bg*) gene of 200 samples were amplified. The prevalence of *G. duodenalis* recorded from patient samples collected in Karonga, Lilongwe and Mangochi were 17.2%, 9.7% and 3.1%, respectively. The overall prevalence of *G. duodenalis* by the *bg* gene was 10% (20/200) (Table 2.5).

The prevalence of *G. duodenalis* by *bg* was significantly different to the prevalence of *G. duodenalis* by 18S SSU-rRNA and *tpi* (Fisher's exact test, $p = 0.001$) and the p-value of the prevalence of *G. duodenalis* in Malawi by *bg* and *gdh* was 0.003, which is also significantly different. However, there were no significantly different between the prevalence of *G. duodenalis* Malawi by *bg* among those from each region in Malawi (Fisher's exact test, $p > 0.05$).

Out of 20 samples from three regions in Malawi, 16 (80%) were successfully sequenced and genotyped. The results of the sequencing analysis matched the GenBank reference (Table 2.7). The prevalence of Assemblage A by *bg* was significantly different to the prevalence of Assemblage AB by *bg* (Fisher's exact test, $p = 0.002$) and the prevalence of Assemblage B by *bg* was significantly different to the prevalence of *G. duodenalis* that was determined by *bg* (Fisher's exact test, $p = 0.005$). However, there were no significant difference between the prevalence of Assemblage A and *G. duodenalis* that was determined by *bg* (Fisher's exact test, $p > 0.05$) (Table 2.7). By using the *bg* gene, 33% (1/3) contained sub-assemblage AI, 67% (2/3) displayed sub-assemblage AII. For Assemblage B, 46% (5/11) contained sub-assemblage B heterogenous, 27% (3/11) showed sub-assemblage BIII and 27% (3/11) belonged to sub-assemble other B.

Table 2.7 *G. duodenalis* genotype by β -giardin in children under 5 year of age from three regions in Malawi

Source	Assemblage A	Assemblage B	Seq. not determined	Total
Karonga	18.2% (2/11)	72.7 % (8/11)	9.1% (1/11)	11
Lilongwe	14.3 % (1/7)	42.9% (3/7)	42.8% (3/7)	7
Mangochi	0%(0/2)	100% (2/2)	0% (0/2)	2
Total	15% (3/20)	65% (13/20)	20% (4/20)	20

2.4.2.4 Analysis of glutamate dehydrogenase (*gdh* gene)

The 530 bp fragment of the *gdh* gene of 200 samples was amplified. The prevalence of *G. duodenalis* recorded from patient samples collected in Karonga, Lilongwe and Mangochi were 21.9%, 20.8%, and 25.5%, respectively. The overall prevalence *G. duodenalis* in Malawi by the *gdh* gene was 21% (42/200) (Table 2.5). The prevalence of *G. duodenalis* in Malawi by *gdh* was significantly different to the prevalence of *G. duodenalis* by *bg* (Fisher's exact test, $p = 0.003$). However, the prevalence of *G. duodenalis* by *gdh* was not significantly different to the prevalence of *G. duodenalis* by 18S SSU-rRNA and *tpi* (Fisher's exact test, $p > 0.05$). There were no significant differences between the prevalence of *G. duodenalis* by *gdh* to those by 18S SSU-rRNA and *tpi* (Fisher's exact test, $p > 0.05$). There were no significant differences between the prevalence of *G. duodenalis* by *gdh* among those from each region in Malawi (Fisher's exact test, $p > 0.05$).

By using the *gdh* gene, 56% (18/31) contained sub-assemblage BIV H43 (EF507682), 38% (12/31) displayed sub-assemblage BIV (KC96064.1) and 6% (2/31) belonged to sub-assemblage B Cla145 (HM134212.1). Out of 42 samples, 31 (73.8%) were successfully sequenced and genotyped. The results of the sequencing analysis matched the GenBank reference (Table 2.8).

Table 2.8 *G. duodenalis* genotyped by *gdh* gene in children under 5 year of age from three regions in Malawi

Source	Assemblage B	Seq. not determined	Total
Karonga	50 % (7/14)	50% (7/14)	14
Mangochi	100% (11/11)	0% (0/11)	11
Lilongwe	76.5% (13/17)	23.5% (4/17)	17
Total	73.8% (31/42)	26.2% (9/42)	42

2.4.3 Genotyping and subtyping of *Cryptosporidium* spp.

2.4.3.1 PCR amplification of SSU-rRNA gene

The 826 bp fragment of the SSU-rRNA gene of 200 samples were amplified. The prevalence of *Cryptosporidium* spp. recorded from samples collected in Karonga, Lilongwe and Mangochi were 6.3%, 22.2%, and 3.1% respectively. The overall prevalence of *Cryptosporidium* spp. in Malawi by SSU-rRNA gene was 11% (22/200) (Table 2.9)

Table 2.9 Detection of *Cryptosporidium* spp. (SSU-rRNA and GP60 genes) in children under 5 year of age from three regions in Malawi

Source	SSU-rRNA	GP60	Mixed infection (<i>G. duodenalis</i> + <i>Crypto</i>) by SSU- rRNA	Total of positive for <i>Crypto</i>
Karonga	6.3% (4/64)	6.3% (4/64)	4.7% (3/64)	6.3% (4/64)
Lilongwe	22.2% (16/72)	13.9% (10/72)	19.4% (14/72)	22.2% (16/72)
Mangochi	3.1% (2/64)	3.1% (2/64)	3.1% (2/64)	3.1% (2/64)
Total	11% (22/200)	8% (16/200)	9.5% (19/200)	11% (22/200)

2.4.3.2 PCR- Restriction Fragment Length Polymorphism

To identify the species and genotypes of *Cryptosporidium*, the secondary products were digested with *SspI* and *VspI* (Figure 2.3). For species diagnosis, *SspI* was used to digest the SSU PCR product. Digestion with *SspI* showed a characteristic pattern of three major visible bands in both *C. parvum* and related species. These include a 449 base pair fragment, a 254 and 111 bp bands (*C. parvum* 'human' genotype (*C. hominis*) type 1) (Figure 2.3, upper gel, Lanes 2, 4, 6, 8 and 9) but for *C. parvum* 'bovine' genotype (*C. parvum*) type 2 include a 449 base pair fragment, a 254 and 108 bp bands (Figure 2.3, upper gel, Lanes 3 and 5). Lanes 1 and 7 showed mixed infection with both genotypes.

For genotyping of *C. parvum*, the different genotypes were identified by *VspI* for digestion. Lanes 2, 4, 6, 8 and 9 showed fragments corresponding to *C. parvum* 'human' genotype (*C. hominis*) of approximately 561 bp (Figure 2.3, lower gel). Lanes 3 and 5 showed fragments identified as *C. parvum* 'bovine' genotype (*C. parvum*) of approximately 628 bp (Figure 2.3). Visualisation of PCR-RFLP of *Cryptosporidium* spp. based on SSU-rRNA gene from Lilongwe samples, Lane L: Ladder (100 bp marker), Lane 2, 4, 6, 8, and 9 = *C. hominis* (H); Lane 1 and 7: Mixed infection with both genotype (M), Lane 3 and 5 = *C. parvum* (P), Lane 10 = +ve (H), Lane 11 = +ve (P), Lane 12 = -ve, *SspI* (upper gel), *VspI* (lower gel). Lanes 1 and 7 showed mixed infection with both genotype.

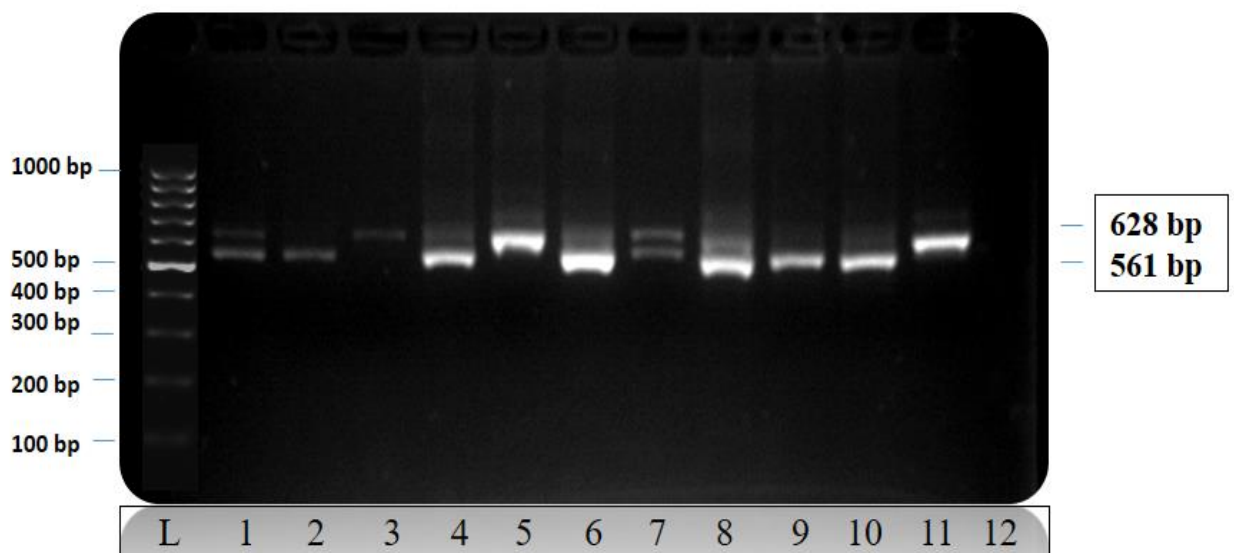
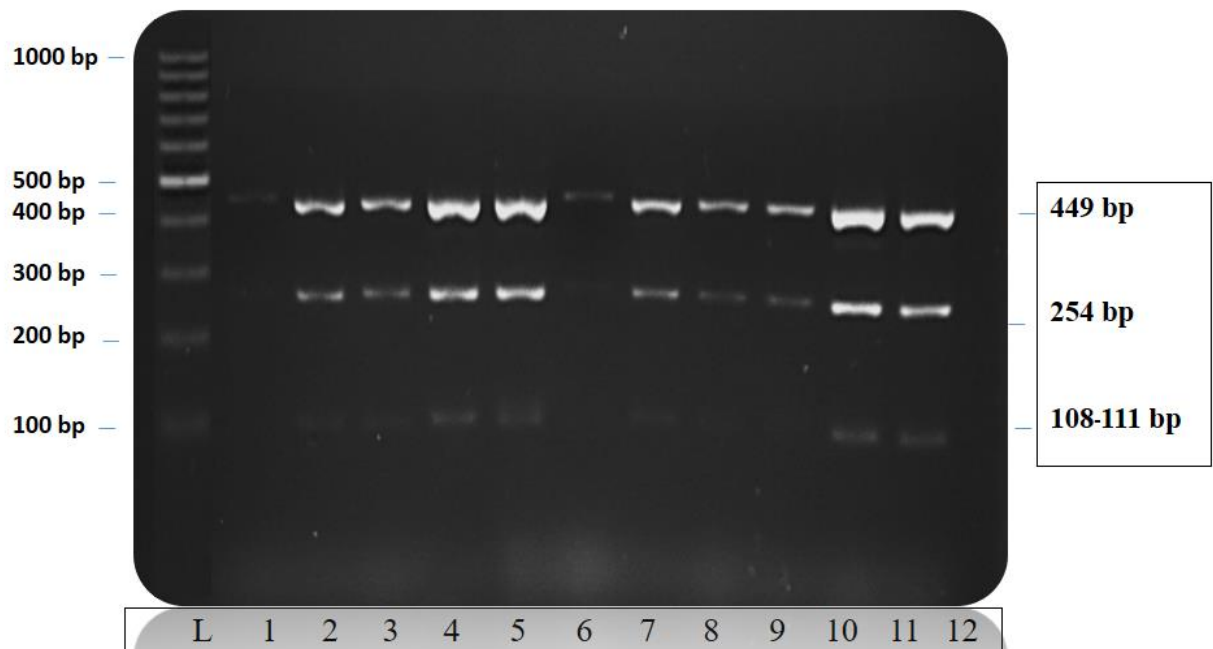


Figure 2.3 Visualisation of PCR-RFLP of *Cryptosporidium* spp. based on SSU-rRNA gene from Lilongwe samples, Lane L: Ladder (100 bp marker), Lane 2, 4, 6, 8, and 9 = *C. hominis*; Lane 1 and 7: Mixed infection with both genotype, Lane 3 and 5 = *C. parvum*, Lane 10 = +ve (*C. hominis*), Lane 11 = +ve (*C. parvum*), Lane 12 = -ve, Notice: *SspI* (upper gel), *VspI* (lower gel).

2.4.3.3 PCR amplification of GP60

The 400-450 bp fragment of the GP60 gene of 200 samples was amplified. The prevalence of *Cryptosporidium* spp., recorded from samples collected in Karonga, Lilongwe and Mangochi were 6.25%, 13.88% and 3.125% respectively. The overall prevalence of *Cryptosporidium* spp. in Malawi by GP60 gene was 8% (16/200) (Table 2.9).

2.4.3.4 SSU rRNA-based PCR-RFLP, GP60 and DNA sequence analysis

Restriction analysis of the SSU PCR products showed that 10 of the 22 PCR-positive samples had *C. hominis*, eight had *C. parvum* and four samples were infected with both species (*C. hominis*/*C. parvum*) (Table 2.10). Sequence analysis of the SSU rRNA and GP60 genes confirmed the species identification by SSU rRNA-based PCR-RFLP analysis (Table 2.10). Mixed infection (*C. hominis*/*C. parvum*) was found in 25% of samples from Lilongwe. Moreover, *C. hominis* and *C. parvum* were found in the same percentage (37.5%). However, only *C. hominis* was found in Mangochi whereas both species were found in the same percentage (50%) in Karonga (Table 2.10).

Table 2.10 Genotyping results by using SSU rRNA-based PCR-RFLP, GP60 and sequencing for *Cryptosporidium* spp. from 3 regions in Malawi

Sources	<i>C. hominis</i>	<i>C. parvum</i>	Mixed infection (<i>C. hominis</i> / <i>C. parvum</i>)
Karonga	50% (2/4)	50% (2/4)	0% (0/4)
Lilongwe	37.5% (6/16)	37.5% (6/16)	25% (4/16)
Mangochi	100% (2/2)	0% (0/2)	0% (0/2)
Total	45.40% (10/22)	36.40% (8/22)	18.20% (4/22)

2.4.4 The correlation between gender, rotavirus co-infection, age with *G. duodenalis* assemblages and *Cryptosporidium* spp.

2.4.4.1 The correlation between gender and *G. duodenalis* assemblages and *Cryptosporidium* spp.

Female patients 29.2% (28/96) were found to be infected with *G. duodenalis* and male patients 26.9% (28/104) were infected with *G. duodenalis* (Figure 2.4). Of the 22 samples from males, 11 (50%) were infected with assemblage A/B, five (22.72%) with assemblage A, and six (27.27%) with assemblage B. Of 20 samples from females, assemblage B was detected in nine (45%), assemblage A in three (15%), and assemblage AB in eight (40 %) (Figure 2.7). More females were found to be infected with *G. duodenalis* than males but not significantly different (Fisher 's exact test, $p > 0.05$). Moreover, there was no significant difference between male and female subjects in the proportion of samples that were positive for the three assemblages (A, B and A/B) (Fisher 's exact test, $p > 0.05$) (Figure 2.7).

For *Cryptosporidium*, females were infected with 18.9% (18/95) while males were infected with 14.7% (14/105) (Figure 2.4). More females were found to be infected with *Cryptosporidium* spp. than males. Furthermore, there were no significant differences between male and female subjects in the proportion of samples that were positive for *Cryptosporidium* spp. (Fisher's exact test).

2.4.4.2 The correlation between rotavirus co-infection and *G. duodenalis* assemblages/*Cryptosporidium* spp.

Among 56 patients infected with *G. duodenalis*, 43 (76.8%) showed no co-infection with rotavirus. The number of patients who were infected with both rotavirus and *G. duodenalis* (23.2%, 13/56) (Figure 2.5). There was a significant difference between the number of subjects who had rotavirus co-infection and those who had no rotavirus co-infection (Fisher 's exact test, $p = 0.008$) (Figure 2.5). The different *Giardia* assemblages were distributed through patients with rotavirus co-infection as follows. The most common assemblage was B (54.54%), then AB (36.4%) and least common was A (9.1%). The *Giardia* assemblage distribution for patients with no rotavirus co-infection were AB (48.4%) then B (29%) and the least common assemblage was A (22.6%). Moreover, there was no

significant difference between the *G. duodenalis* infected patients who had rotavirus co-infection and those who didn't in the proportion of samples testing positive for the three assemblages (A, B and AB mixed) (Fisher's exact test) (Figure 2.7). For those patients infected with *Cryptosporidium* spp. The results were as follows; Patients with no rotavirus co-infection numbered (81.8%, 27/33), patients with rotavirus co-infection numbered (18.2%, 6/33) (Figure 2.5). The much higher number of patients with no rotavirus co-infection demonstrates that there is no significant association with rotavirus. There was a significant difference between subjects who had *Cryptosporidium* spp. with rotavirus co-infection and those who had *Cryptosporidium* spp. without rotavirus co-infection (Fisher's exact test, $p = 0.001$).

2.4.4.3 The correlation between age and *G. duodenalis* assemblages and *Cryptosporidium* spp.

Regarding age, patients were divided into 3 groups; less than 1 year, 1 to 2 years, and 2 to 4 years. Children infected with *G. duodenalis* were most commonly identified in the age group 1 to 2 years (40.9%, 9/22), followed by the age less than 1 year (36.4%, 43/118) and the age group 2 to 4 years (30.8%, 4/13) (Figure 2.6). Otherwise, Assemblage AB (55.56%) was detected more frequently than assemblage A (22.22%) and assemblage B (22.22%) in the age group less than 1 year whereas Assemblage B (70%) was detected more frequently than assemblage A (10%) and assemblage AB (20%) in the age group 1 to 2 years. For the age group 2 to 4 years found assemblages B and AB at the same percentage (40%) and the rest was assemblage A (20%) (Figure 2.7).

For *Cryptosporidium* spp., children were divided 3 groups; less than 1 year, 1 to 2 years, and 2 to 4 years. Children infected with *Cryptosporidium* spp. were in the age from 2 years to 4 years (30.8%, 4/13), followed by the age less than 1 year (25.5%, 24/94) and the age group 1 to 2 years (15.4%, 4/26) (Figure 2.6).

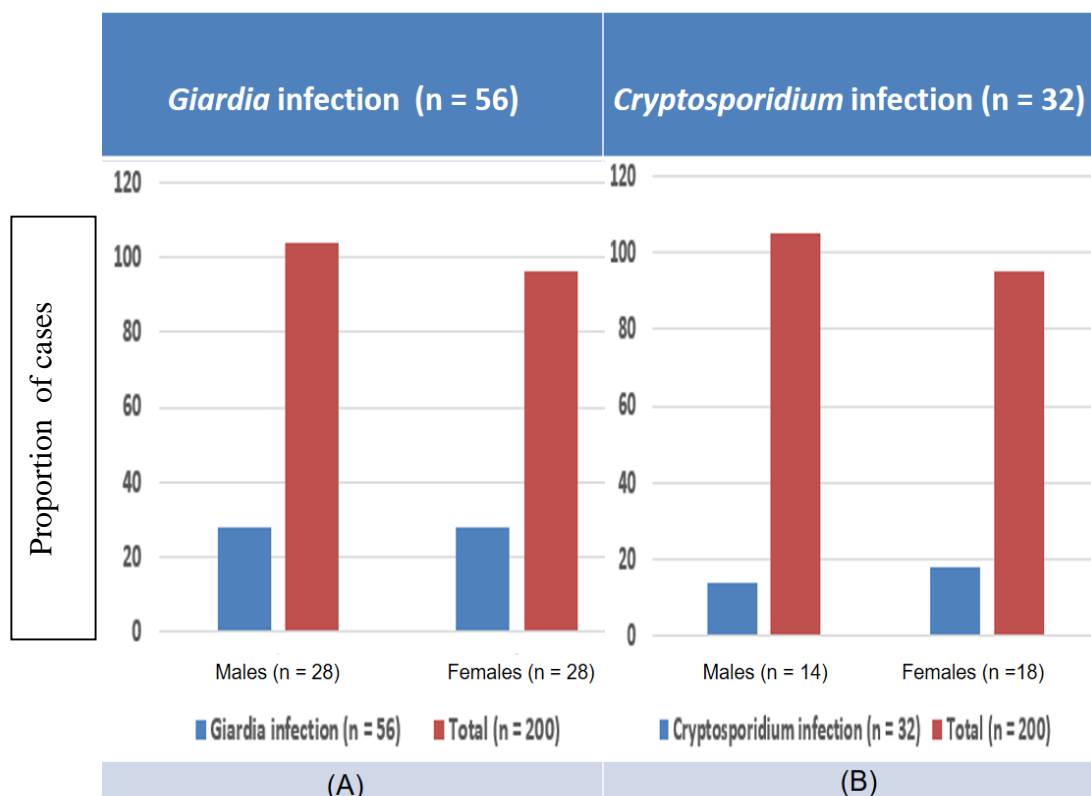


Figure 2.4 The number of positive samples for *G. duodenalis* (A) and *Cryptosporidium* (B) in male and female children

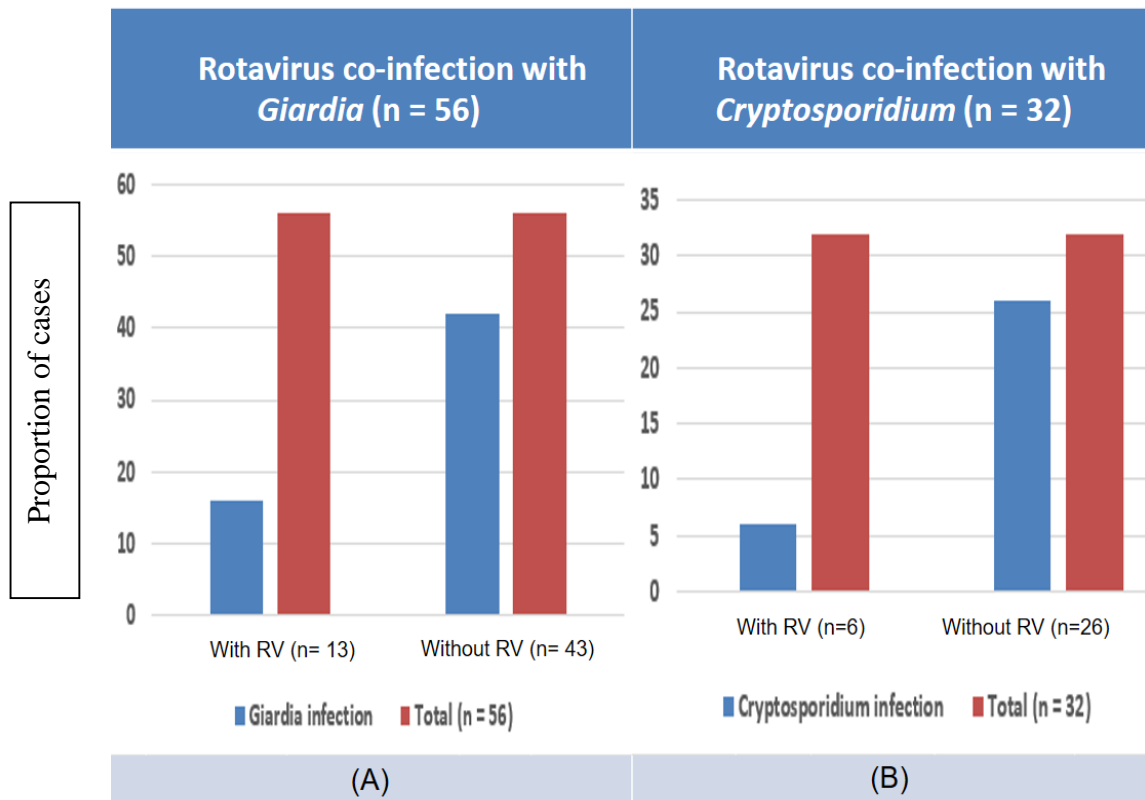


Figure 2.5 The number of positive samples for *G. duodenalis* (A) and *Cryptosporidium* (B) in two groups of children with or without rotavirus infection

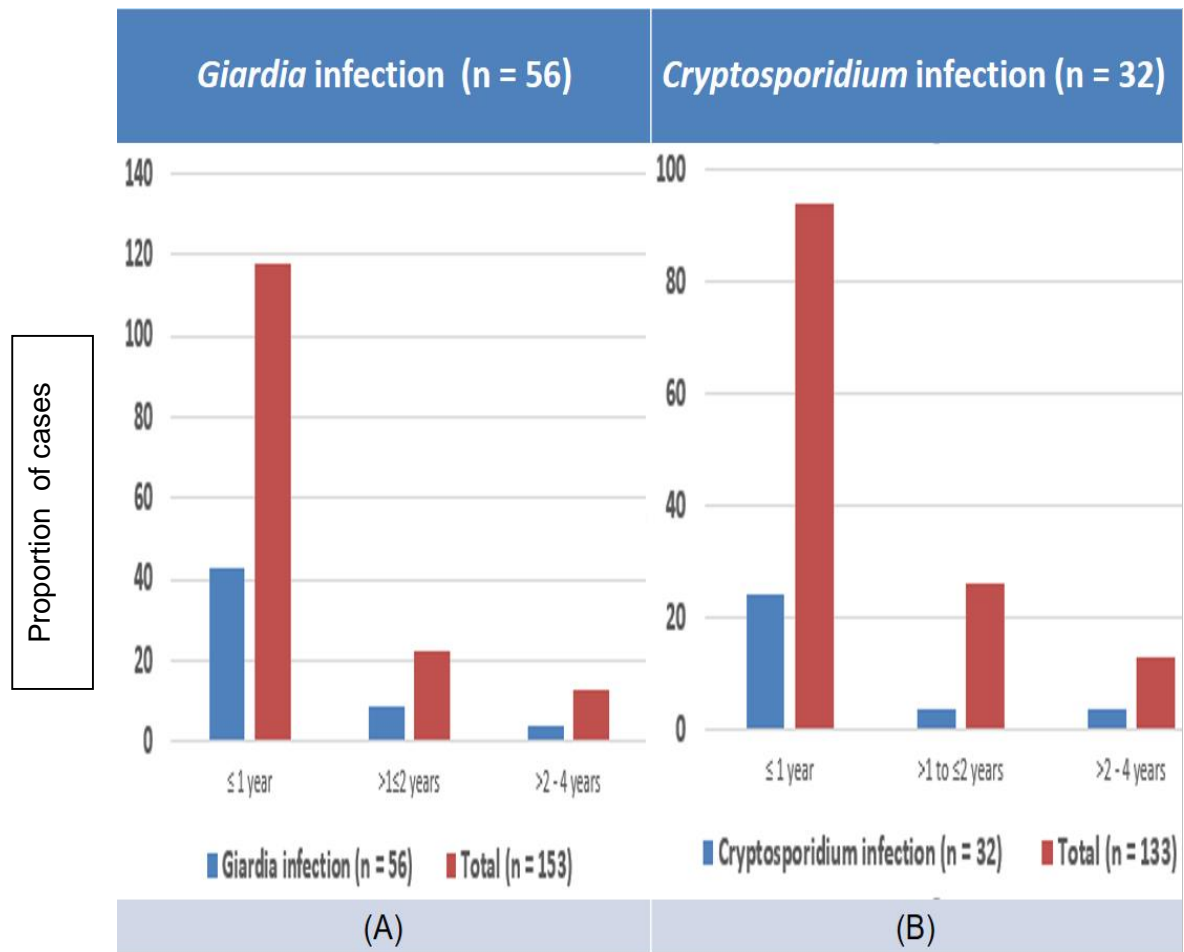


Figure 2.6 The number of positive samples for *G. duodenalis* (A) and

***Cryptosporidium* (B) in different age groups.**

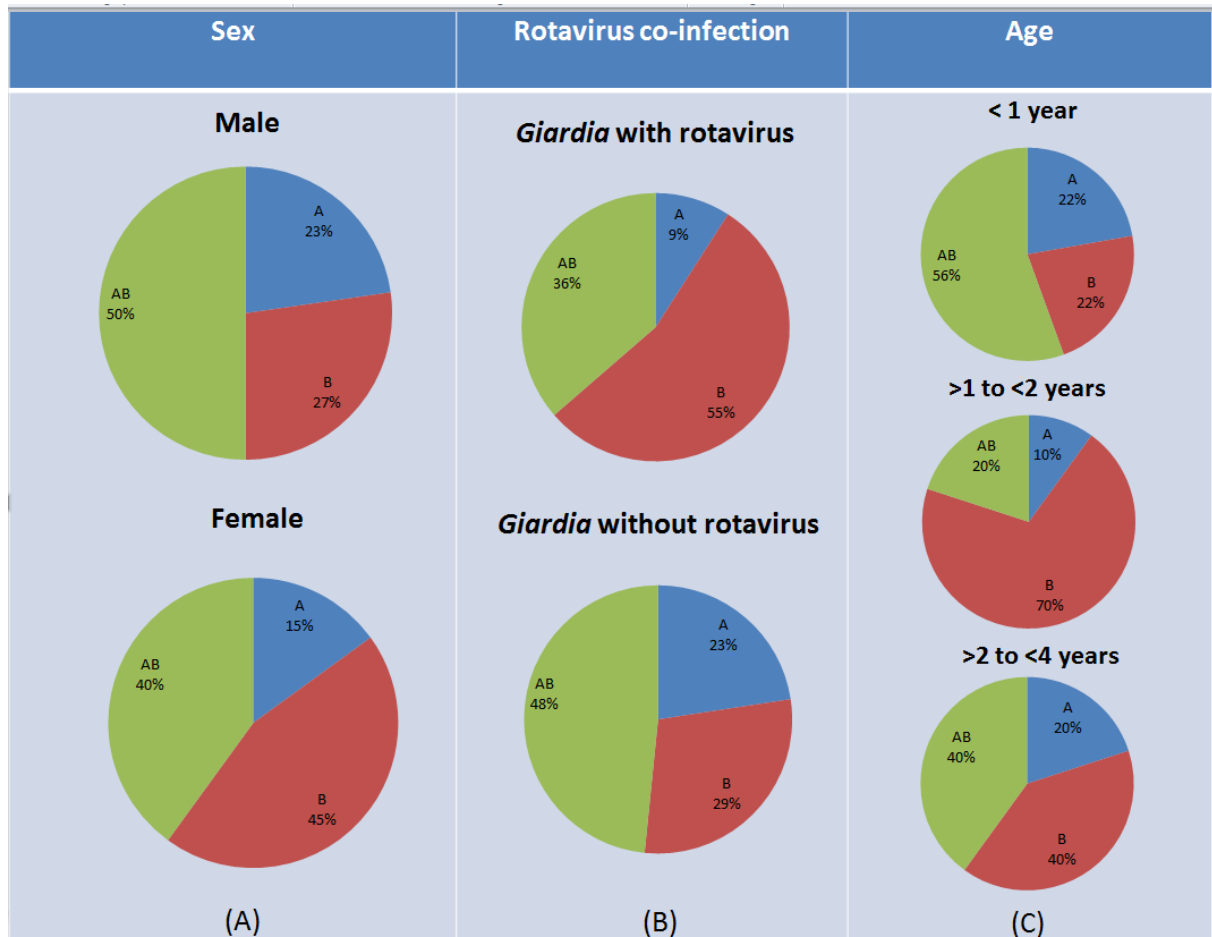


Figure 2.7 The percentage of *G. duodenalis* assemblages/ found in male and female patients (A), the percentage of *G. duodenalis* assemblages found in two groups of children with or without rotavirus infection (B), and the percentage of *G. duodenalis* assemblages found in different age groups (C).

2.4.5 Age distribution of *G. duodenalis* and *Cryptosporidium* spp. cases in Malawian

Children

As mentioned above, most of the cases occurred in children younger than 1 year. However, children had a peak occurrence of Giardiasis at 7 and 8 months of age whereas only one or two cases were found in children older than 1 year (the age from 13 months to 26 months and at the age 38 months) (Figure 2.8). Otherwise, children at 1 month and 4 months of age were infected by *G. duodenalis* in one and two cases, respectively. However, there were none found children infected *Giardia* at 2 and 3 months of age.

The occurrence of *Cryptosporidium* spp. was particularly high in children less than 1 year of age, especially, the 6-12 months of age. However, children had a peak occurrence of Cryptosporidiosis at 6 and 9 months of age, four cases were found in both ages whereas only one case was found in children at 1, 2, 4, and 5 months of age and one case was also found in children older than 1 year (at the age 13, 15, 22, 24, 25 and 38 months). Otherwise, there were none found children infected *Cryptosporidium* spp. at 3 months and 16-21 months of age (Figure 2.8).

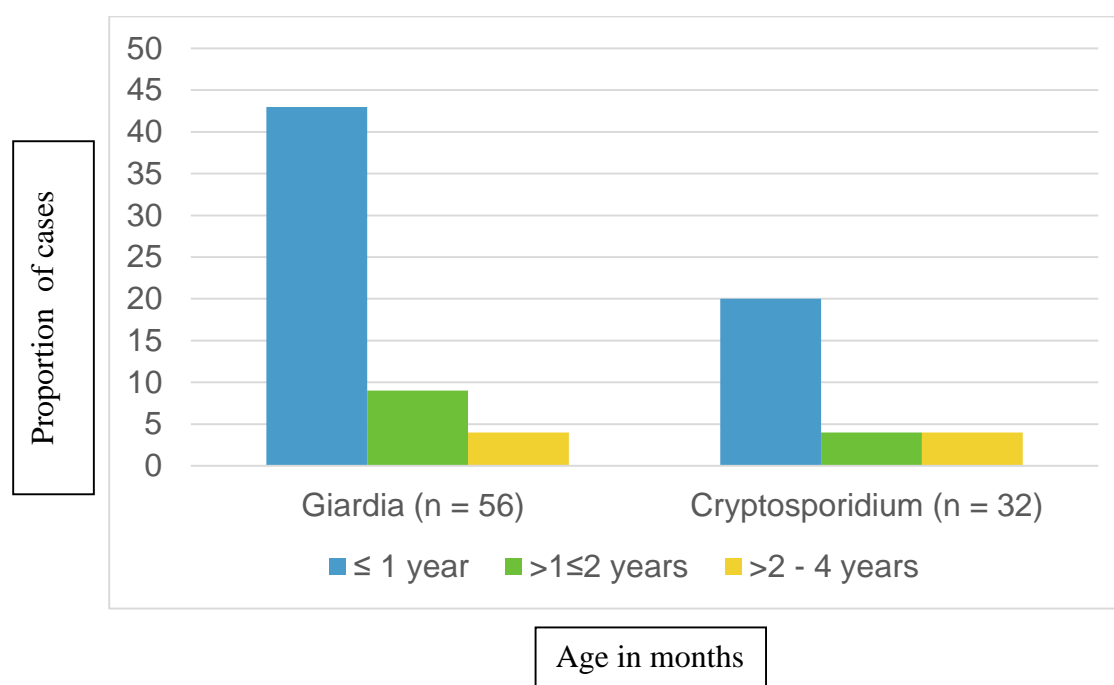


Figure 2.8 Age distribution of *G. duodenalis* (n=56) and *Cryptosporidium* spp.

(n=32) cases in Malawian Children

2.4.6 Seasonal distribution of *G. duodenalis* and *Cryptosporidium* spp. in Malawian Children

There are two seasons in Malawi. The dry season lasts from April through November and the wet season is from December to March. Even in the wet season, the rains are usually short-lived storms. The distribution of the *G. duodenalis* and *Cryptosporidium* spp. cases peaked between June and October. These months approximately corresponded to the dry season in this region. Conversely, during the rainy season months (December-March), the number of cases dropped for both parasites. However, children who had no *Cryptosporidium* spp. infection were found in February, April and June whereas children who had no *G. duodenalis* infection were found in April only (Figure 2.9).

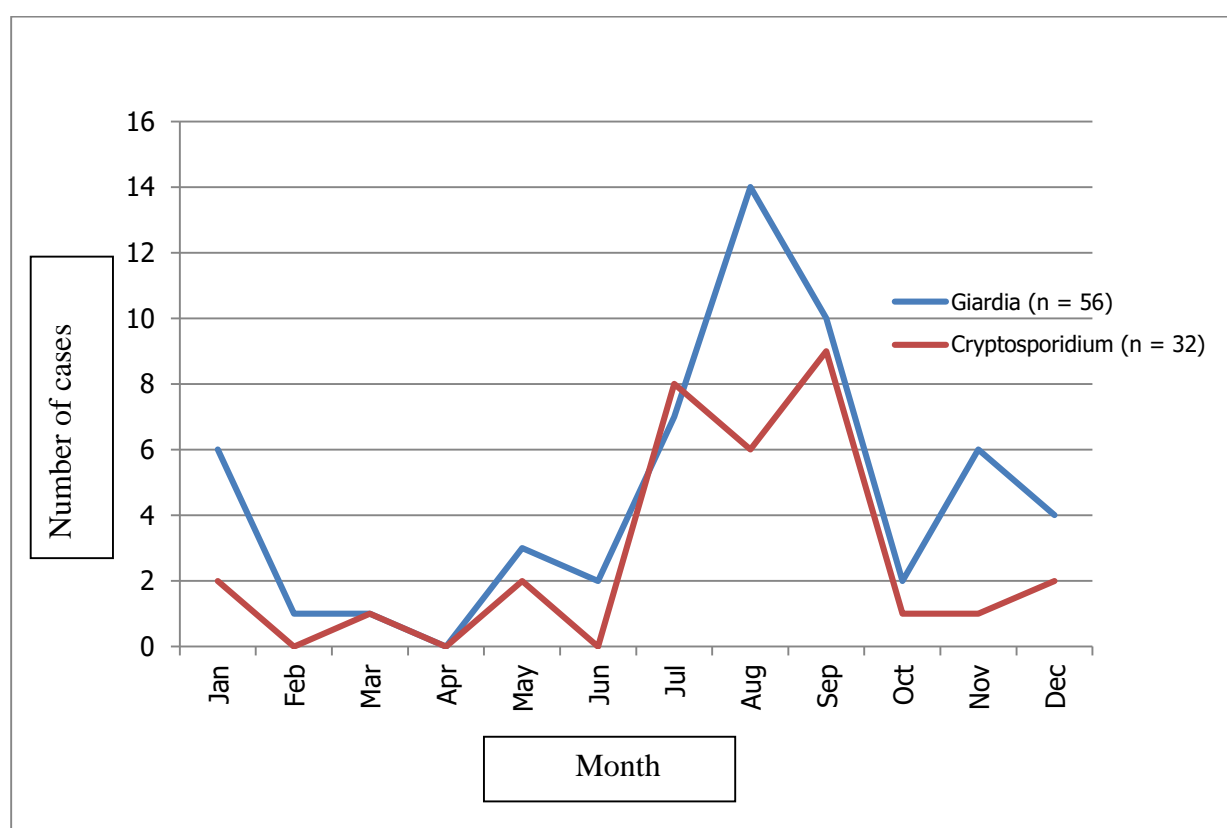


Figure 2.9 Seasonal distribution by month of *G. duodenalis* (n=56) and *Cryptosporidium* spp. (n=32) cases in Malawian Children

2.5 DISCUSSION

The faecal samples analysed in this Chapter were recovered from children with diarrhoea in Malawi, under 5 years of age, living in different geographical regions. Faecal samples were screened for both *Cryptosporidium* and *Giardia* microscopically using the Direct Fluorescence Assay (DFA) and parasite isolates were typed by using multilocus genotyping, Nested-PCR and/or sequencing of SSU-rRNA, *tpi*, *bg* and *gdh* genes for *Giardia*. For genotyping of *Cryptosporidium* spp., DNA sequence analysis of the 60 kDa glycoprotein (GP60) gene was used for subtyping of *C. parvum* and *C. hominis*.

The results demonstrated that DFA was more sensitive than PCR for detecting *Cryptosporidium* whereas the opposite was observed for *Giardia*. The samples were kept frozen for a long time and they were thawed many times. In general, the storage of specimens in the freezer (at -80 degree celcius) with no preservative gave good results for DNA extraction but freezing and thawing many times could degrade the DNA. Jongwutiwes *et al.* (2002) recommended the use of 75% alcohol as co-pro-preservative for the long term storage in the case of *Cryptosporidium* research (Jongwutiwes *et al.*, 2002). *Giardia* cysts are more delicate than *Cryptosporidium* oocysts and are easily broken, so they may not be detectable by DFA after freezing/thawing of faeces (but DNA is released more easily). On the other hand, *Cryptosporidium* oocysts have a thicker wall and they may remain intact for DFA detection (whereas DNA is more difficult to extract, leading to a lower detection rate by PCR compared to *Giardia*).

It is difficult to determine whether the faecal samples that were diagnosed positive by DFA but not by PCR were false negatives on PCR, or false positives on DFA. However, it should be recorded that the number of oocysts in the DFA positive, PCR negative samples were very low (<5 per field of view). Therefore, other structures might have been mistaken for oocysts when no counter stain was used. The DFA test detects only intact *Giardia* cysts or *Cryptosporidium* oocysts. As a result of thawing and freezing several times, this could have resulted in some DNA degradation, and made it impossible to remove the PCR inhibitors from the samples by washing before DNA extraction. The DFA technique can be helpful in determining *Cryptosporidium* spp. infection and it provides a valuable tool for the rapid, specific and sensitive detection of *Cryptosporidium* in faecal samples.

Our results show that the screening assays used here in the detection of *Giardia* and *Cryptosporidium* infection had variable sensitivities. PCR generally identified a higher number of *Giardia* infections than the microscopy based DFA assay when testing the same set of samples. The reverse relationship was observed with *Cryptosporidium*. The overall prevalence of *Cryptosporidium* spp. by DFA and PCR (SSU-rRNA) was 23.48% (31/132) and 11%(22/200), respectively. Thus the DFA test is more sensitive than PCR. This result corresponds to a report from New York State, in which faecal samples were designated positive for *Cryptosporidium* on DFA 7.4% (13/175) of cases, whereas on SSU-rRNA based PCR, only 5.1% (9/175) of cases were positive (Burton *et al.*, 2010). The differences in assay sensitivity are perhaps due to the specificity of the DFA reagents and/or in extraction variation between the *Giardia* trophozoite/*Cryptosporidium* oocyst stages.

This study demonstrated the prevalence of *Giardia* by *tpi* gene, assemblages A (14.29%), assemblage B (26.79%) and mixed infection (A/B mixed, 33.93%) that were found in symptomatic patients in Malawi. From a previous study, assemblage A and assemblage B have been reported as the most common genotypes found in humans (Sprong *et al.*, 2009). Many studies found that the Assemblages A and B have ratios in various proportions. Several reports demonstrate that the assemblage B was found to be predominant in some countries. In Malaysia, the assemblage B was found in 98% of isolates and the assemblage A was detected in 2% of isolates. Sahagun *et al.* (2008) studied 108 Spanish patients; assemblage A (40%), assemblage B (56%), and mixed infection of both assemblages (4%) was detected (Sahagun *et al.*, 2008). Amar *et al.* (2002) who studied *Giardia* in human patients with symptomatic giardiasis by using PCR-RFLP analysis and the *tpi* gene found 27% contained assemblage A2, 64% displayed assemblage B, and 9% showed a mixture of assemblage A2 and B. It was concluded that assemblage B was predominantly followed by assemblage A and mixed infection respectively (Amar *et al.*, 2002).

In contrast, some researchers found that assemblage A was predominant. A study by Caccio *et al.* (2002) showed that the assemblage A (80%, 24/30) and the assemblage B (20%, 6/30) were found in patients in Italy by using SSU-rRNA and PCR-RFLP (*bg* marker) (Cacciò *et al.*, 2002). Moreover, in Peru, it was also found the assemblage A (62.5%, 10/16) more than the assemblage B (37.5%, 6/16) from 16 samples that were characterised (Perez Cordon *et al.*, 2008). Furthermore, in some

countries in South America such as Brazil and Mexico found all samples were identified as the assemblage A by using *bg* for PCR marker (Lalle *et al.*, 2005; Volotao *et al.*, 2007).

In the present study, PCR was observed to have the highest amplification rate at the *tpi* locus (28%, 56/200), followed by the *gdh* locus (21%, 42/200) and at the *bg* locus (10%, 20/200), respectively. Reasons why PCR amplification of the *bg* gene resulted in a lower detection rate than the *tpi* and *gdh* genes may be due to the fact that the genetic variation within the *bg* gene is higher than within the *tpi*, *gdh* and highly conserved SSU genes. Other factors potentially influencing PCR efficiency were consistent between the different assays, such as the quality and quantity of DNA template and the quality of the DNA polymerase used.

The differential amplification success we observed at the MLST loci has often been reported in studies using the same primers. The amplification success rate reported by a study of symptomatic patients from Germany was 92.4 % for the *bg*, 44 % for the *tpi* and only 25.7 % for the *gdh* locus (Broglia *et al.*, 2013). It has been suggested that the lack of amplification of certain loci could be due to the presence of nucleotide mismatches between the PCR primers and the genomic sequences (Broglia *et al.*, 2013), leading to certain primers being unable to amplify particular sub-genotypes as suggested for the *bg* primers in respect of particular assemblage B sub-types (Robertson *et al.*, 2007).

Primers are designed to bind 'conserved' regions in the amplified genes; excessive mismatches in the binding regions of primer sequences could potentially result in the failure of PCR amplification for some isolates of *G. duodenalis*. However, numerous molecular data have confirmed both inter- and intra-assemblage genetic variations (Feng & Xiao, 2011). Variable base substitution rates of *G. duodenalis* have been reported at different genetic loci (Wielinga & Thompson, 2007). Such differences in nucleotide sequences, especially base variations in primer binding sites, might lead directly to the failure in the primer-template binding in some *G. duodenalis* isolates. PCR analysis by Scorza *et al.* (2013) showed amplification of the *gdh*, *bg*, and *tpi* genes in 91.0% (172/189), 84.8% (145/171) and 19.8% (34/172) of *G. duodenalis* cyst-positive samples from the mammalian faeces, respectively (Scorza *et al.*, 2013). PCR amplification rates were 100% (63/63), 61.9% (39/63) and 56.0% (14/25), respectively at the *gdh*, *bg* and *tpi* loci in a study of molecular characterization of *Giardia* isolates from clinical infections following a waterborne outbreak (Robertson *et al.*, 2007).

In another comparative study of the *tpi* and *gdh* genes for detection and genotyping of human-derived *G. duodenalis* isolates, the *tpi* gene was amplified from 96.2% (25/26) of samples, whereas only 81% (21/26) were positive when the *gdh* gene was targeted (Bertrand *et al.*, 2005). It was also observed that PCR amplification rates differed between different assemblages. In the main, whichever gene was amplified, assemblage A had a higher amplification rate than assemblage B. This result might be related to the fact that assemblage B exhibits higher allelic sequence heterogeneity and genetic recombination compared to assemblage A (Kosuwin *et al.*, 2010; Teodorovic *et al.*, 2007). Therefore, the reasons for different PCR amplification rates of different assemblages at different loci need to be better understood by a more intense, systematic study and with more complete genetic data from a large number of *G. duodenalis* isolates in the future. The present study is the first report describing genetic characterisations of the *tpi*, *bg* and *gdh* genes of *G. duodenalis* isolates in children under 5 years of age with diarrhoea, living in diverse geographical regions in Malawi. In order to have a greater understanding of how *G. duodenalis* interacts between host and pathogen and the degree to which animals can cause Giardiasis in humans, further epidemiological studies should be conducted where humans and animals share a close proximity and/or in locations where infection is endemic.

The molecular analysis of assemblage B isolates was made more difficult by the high frequency of sequences occurring with heterogeneous positions that had overlapping nucleotides which were noticed at an average of 20% of the sequences across the three loci. A small number of assemblage B isolates demonstrating heterogeneous templates could depict mixed sub-assemblage infections but the identification could not be confirmed beyond doubt. Heterogeneous sequences occur regularly in assemblage B isolates (Sprong *et al.*, 2009; Trout *et al.*, 2004), and their presence has been seen following DNA extraction and sequencing even from single *Giardia* cysts and between different cysts isolated from the same patient (Ankarklev *et al.*, 2012). As well as mixed sub-assemblage infections, the presence of heterogeneous sequences could be due to the presence of two nuclei in *Giardia*, which are believed to acquire mutations and separately evolve thus leading to allele sequence heterozygosity (e.g. the appearance of nucleotidic differences in the sequence between alleles of the same gene) (Ryan *et al.*, 2013). The overall level of allele sequence heterozygosity has been estimated to be significantly higher in the genome of *Giardia* assemblage B compared with assemblage A (Franzén *et al.*, 2009); this may help explain the higher occurrence

of heterogeneous sequences observed in assemblage A as when compared to assemblage B. The allocation to a particular sub-assemblage of isolates was not immediate, and this was explained by discrepancies between the different markers. A large number of positions with allelic sequence variation suggests that gene sequences could be different between the two nuclei in the cell of *Giardia*. Analysis of the ratio between the major and minor nucleotide in the variable positions show that ratios that deviate from 1:1 are common, which suggests that the two chromosomes within a single nucleus may be different.

My hypotheses about genetic diversity from the different regions in Malawi were that: 1) there will be anthroponotic species present in urban areas. 2) there will be zoonotic species in rural areas. 3) there will be mixed infection (anthroponotic and zoonotic species) in both urban and rural areas. The predominance of Assemblage B (27.27%, 6/22) was found in urban Lilongwe whereas the Assemblage A was detected in 13.6% (3/22). It indicates that the anthroponotic route plays a major role in transmission in urban areas. Otherwise, the predominance of Assemblage A (26.32%, 5/19) was found in Mangochi while the Assemblage B was detected in 15.79% (3/19). It demonstrates that the zoonotic route plays a major role in transmission in rural area. Due to the fact that samples were collected from patients who were admitted in the paediatric wards (inpatients) at the Mangochi District Hospital in rural Mangochi. Nonetheless, mixed infection (anthroponotic and zoonotic species) was found in both urban and rural areas. In Lilongwe, the prevalence of A/B mixed was found to be 45.45% (10/22) which was more common than both assemblage A (13.6%, 3/22) and assemblage B (27.27%, 6/22). It is possible that because the samples were collected from a big, hospital serving a large area (Kamuzu Central Hospital), there could be patients from both countryside (rural areas) and urbanized (urban area) presenting themselves or being admitted to this hospital. Undoubtedly, *Giardia* both anthroponotic and/or zoonotic species could be found in Lilongwe. Similarly, in Mangochi, the prevalence of A/B mixed was found in 31.58% (6/19) which were found more than assemblage A (26.32%, 5/19) and assemblage B (15.79%, 3/19). Although, it is a medium hospital, there maybe patients from a variety of areas admitted to the paediatric wards (inpatients) at the Mangochi District Hospital in rural Mangochi. So, there might be some chance that we could find both anthroponotic and zoonotic species. In Karonga which has a small rural hospital, namely the Chilumba Rural Hospital. We should have found assemblage A more than the

assemblage B, however, we have not found the assemblage A alone (0/15) whereas the assemblage B and A/B mixed infection were found in 40% (6/15) and 20% (3/15), respectively. This could be due to the patients having had a history of travel to different urban areas where they could have become infected with anthroponotic *Giardia* assemblages. Therefore, when the patients presented themselves to the clinic (outpatients) at a small rural hospital in Karonga, we could find the assemblage B and AB mixed infection but we could not find the assemblage A alone.

With reference to *Cryptosporidium* spp. It is thought that *C. hominis* is spread only between humans and that the main source of infection with *C. parvum* is cattle and other species of farmed livestock. Direct contact with cattle together with indirect transmission in drinking water are major transmission pathways.

From my results, the prevalence of anthroponotic species (*C. hominis*) is equal in number to zoonotic species (*C. parvum*) (37.5%, 6/16) was found in Lilongwe whereas mixed infection was detected in 25% (4/16). It indicates that both anthroponotic (*C. hominis*) and zoonotic (*C. parvum*) are spread in humans in urban areas. Even though, the samples were collected from patients who were admitted at the Kamuzu Central Hospital in urban Lilongwe. Otherwise, the prevalence of anthroponotic species (*C. hominis*) equal to zoonotic species (*C. parvum*) (50%, 2/4) was found in Karonga whereas mixed infection was not detected. It demonstrates that both anthroponotic and zoonotic infections are also spread amongst humans in rural areas. Due to the samples being collected from patients who admitted in a small rural hospital, namely the Chilumba Rural Hospital, Karonga. In contrast, only *C. hominis* was found in Mangochi (100%, 2/2). However, the sample size was small. Although, the Mangochi District Hospital is a medium sized rural hospital, the patients may receive the anthroponotic species from visiting different areas along with the possibility of indirect transmission through drinking water. So there may be a chance to become infected with *C. hominis* from other urban areas. The overall prevalence of *C. hominis* (45.4%, 10/22) and *C. parvum* (36.4%, 8/22) infections were not significantly different. So, this result could not support the hypothesis that *C. hominis* was spread only between humans as we could demonstrate *C. parvum* was also spread between humans in both rural and urban areas. This result was different from other published papers in that they found a higher *C. parvum* infection rates in rural more than in urban areas (Morse *et al.*, 2007) whereas our results found the prevalence of *C. hominis* and *C. parvum* infection

rates are equal. Not only *C. parvum* can be found in humans but many of *C. canis* and *C. felis* infection in humans can be found also. Moreover, the transmission of *C. canis* and *C. felis* is through the anthroponotic rather than the zoonotic pathway (Cama *et al.*, 2006). Molecular epidemiological research needs to be carried out in order to completely understand the transmission methodology for *Giardia duodenalis* and *Cryptosporidium* spp. This will help towards a full understanding of the frequencies and significance of zoonotic transmission.

In this study, 29 children had both *Giardia* and *Cryptosporidium*. Other reports have documented co-infection of *Giardia*, *Cryptosporidium* and other parasites. Abdel-Hafeez *et al.* (2012) studied immunosuppressed and immunocompetent patients in Egypt by using the microscopic technique for identification. Co-infection between *C. parvum* (60.2%), *G. duodenalis* (17.6%), *Blastocystis hominis* (12.1%), *Isospora belli* (9.7%) and *Cyclospora caytenensis* (7.8%) were found. This study demonstrated that *C. parvum* had the highest infection rates in the immunosuppressed group and it is indicated that *C. parvum* was the most common in immunosuppressed children (Abdel-Hafeez *et al.* 2012). In Nairobi, Kenya, *Cryptosporidium* spp. (30.5%) and *G. duodenalis* (16%) were found in the same persons (co-infection) in young patients under 5 years of age with diarrhoea detected by Microscopy (Mbae *et al.*, 2013). On the other hand, there was no *Cryptosporidium* and *G. duodenalis* in the same samples (no co-infection) in Western Uganda (Salzer *et al.*, 2012).

Restriction analysis of the SSU PCR products showed that 45.40% PCR-positive samples had *C. hominis*, 36.40% had *C. parvum* and 18.20% infected both species (*C. hominis/C. parvum*). Sequence analysis of the SSU rRNA and GP60 genes confirmed the species identification by SSU rRNA-based PCR-RFLP analysis. This result corresponds to a report from Malawi, 41 of the 43 PCR positive samples had *C. hominis* and two had *C. parvum* (Peng *et al.*, 2003). However, Morse *et al.* (2007) found *C. hominis*, *C. parvum*, *C. hominis/C. parvum*, *C. meleagridis* and *C. andersoni* in children in Malawi. Peng *et al.* (2003) studied the molecular epidemiology of cryptosporidiosis in infant patients in Malawi by using multilocus genotyping analyses of the SSU rRNA, 70 kDa heat shock protein (HSP70) and 60 kDa glycoprotein (GP60) genes. From 69 DNA samples positive by microscopy, 43 positive samples, RFLP-PCR products of the SSU rRNA, sequence analysis of the HSP70 and GP60 gene demonstrated that had *C. hominis* in 41 samples (95.35%, 41/43) and *C. parvum* in two samples (4.65%, 2/43).

Morse *et al.* (2007) investigated the incidence of *Cryptosporidium* spp. infection in paediatric patients in Malawi; oocysts were found in 50 of 848 samples (5.9%) and 43 of them were amplified by PCR-RFLP of 18S rRNA and COWP gene, demonstrating the presence of four species; *C. hominis*, *C. parvum*, *C. meleagridis* and *C. andersoni* and also mixed infection (*C. hominis*/*C. parvum*) was found. In contrast, a study on identification and characterisation of *Cryptosporidium* spp. in Ethiopia found 95.12% were characterised as *C. parvum* the most (39/41), 2.44% (1/41) with *C. hominis* and 2.44% (1/41) with mixed infection. These results indicated that *C. parvum* is the major cause of human cryptosporidiosis (Adamu *et al.*, 2010). In Egypt, Abd *et al.* (2012) detected and characterised *Cryptosporidium* spp. in human patients in Cairo. Only 15 samples were characterised, nine samples (60%) were *C. hominis*, three samples were *C. parvum* (20%) and three samples were mixed infection (*C. hominis*/*C. parvum*, 20%) (Abd El Kader *et al.*, 2012).

Associations between infection and gender, age and *Rotavirus* co-infection for both parasites were analysed from clinical information. This result demonstrated that males tended to be more commonly infected with *Giardia*, whereas the opposite was observed for *Cryptosporidium*. However, there were no significant difference between the proportion of males and females positive for these parasites ($p > 0.05$). No relationship between gender and *Giardia* assemblage infection was found. Only a few studies that have investigated the correlation between gender and *Giardia* assemblages have been determined. This study agreed with previous reports (Breathnach *et al.*, 2010; Lebbad *et al.*, 2011). It corresponds to Minetti *et al.* (2014). There was no difference in the prevalence of the three assemblages (A, B and A/B mixed) between males and females (Minetti *et al.*, 2014; Weerapol, 2010). The occurrence of *Cryptosporidium* was particularly high in children less than 1 year of age. Children less than age 1 year were infected significantly more with *Cryptosporidium* than with *Giardia* ($p < 0.05$). The presence of the parasites was not significantly associated with co-infection with Rotavirus. In addition, the presence of rotavirus infection had no correlation to the *Giardia* assemblages detected in this study.

For this study of the age distribution of *Cryptosporidium* infection, there was a peak in children aged less than 1 year. This bears similarities to a previous study of *Cryptosporidium* infections in

Mexican children. The study found that most cases of *C. parvum* occurred in infants less than one year of age. *C. parvum* was also found to be more prevalent amongst malnourished children ($p < 0.05$) and those infants less than six months of age which were not breast fed ($p < 0.01$) (Javier Enriquez *et al.*, 1997).

With reference to *Giardia*, another study found that in children > 1 year, *G. duodenalis* infection was noticeably less frequent in children who were breast fed (1.9%, 7/368) than those that weren't breast fed (6.5%, 7/107; $p=0.012$). It was also noted that of the children who were > 1 year, those that were breastfed showed a significant association with lower prevalences of *G. duodenalis* infection, 0.4% (1/225), versus 6.0% (3/50; $p = 0.003$). However, as the children aged so the prevalence of *G. duodenalis* increased, breast feeding also appeared to protect the children from *G. duodenalis*. The following study details important information about the prevalence and clinical characteristics of these same intestinal parasites in Tanzanian children (Tellevik *et al.*, 2015). However, Mahmud *et al.* (2001) found that breast feeding prevents both symptomatic and asymptomatic *G. duodenalis* infection among children < 12 months in Egypt (Mahmud *et al.*, 2001). Ignatius *et al.* (2012) also found that infants fed breast-milk appeared to be significantly better protected from infection in a study of Rwandan children, but did not make any connection as to symptomatic or asymptomatic infection (Ignatius *et al.*, 2012). Elsewhere, researchers found a significant predominance of children infected with *Giardia* at the age of less than 5 years and adults who were aged 20-30 years (Bern *et al.*, 2000). In endemic areas the reasons for these two patterns is unclear. The decline in symptoms with age and the fact that most cases occur to those aged under 5 years could be explained by the interplay of transmission intensity, behaviour and immunity. *G. duodenalis* immunity tends to reduce the severity of the symptoms rather than stopping re-infection (Kohli *et al.*, 2008; Solaymani-Mohammadi and Singer, 2010). However, if we accept that ageing appears to improve immunity, it should be considered that factors such as reduced exposure, more hygienic practices/activities, altered diet and improved water supply (maybe at places of learning or work) and other changes in behavioural factors may have a significant affect. (Heimer *et al.*, 2015). In this study, children had a peak occurrence of giardiasis at 7 and 8 months of age. For the occurrence of cryptosporidiosis, it had a peak at 6 and 9 months of age. It is different to the results of the study in Mexico which showed a peak occurrence of cryptosporidiosis at 10 months of age and much

earlier than those reported in other areas (Javier Enriquez *et al.*, 1997). The differences may be explained by the fact that the sample size in this study was small and not truly representative. Other variables such as transmission modes, exposure frequency, immunity growth, geography and sanitation standards, should also be considered.

It should be noted that amongst the characteristics that affected parasitic infection there are three important factors. Immunity of the hosts, this is variable and can depend on items such as the age, behavior, susceptibility and genetic background of the host. Environmental factors which can include local cultures, hygiene and sanitation standards, consumption and preparation methods of local cuisine, health and wellbeing of the local community and geographical and climatic conditions. Finally, factors relating to the parasite itself are important and the following should be taken into consideration; mutation, adaptation to anti-parasitic treatments and natural evolution depending on geographical location.

The seasonal distribution of the cases peaked between June and October. These months approximately corresponded to the dry season in this region. Conversely, during the months of the rainy season, the number of cases dropped. This shows differences to the results of the study on the seasonal transmission of *Cryptosporidium* infection and its association with rainfall in the Queen Elizabeth Central Hospital (QECH), Blantyre, Malawi between August 1997 and March 1999 (Peng *et al.*, 2003) and it is also different to the results in Bangladesh, Kuwait, Guatemala, Mexico and Tanzania as *Cryptosporidium* was most common in the rainy season in these countries. (Bhattacharya & Mandal, 1997; Bern *et al.*, 2000; Iqbal *et al.*, 2001; Javier Enriquez *et al.*, 1997; Tellevik *et al.*, 2015). The prevalence of *Cryptosporidium* in Tanzania was higher in the rainy months (12.9%, 105/812) than in the dry months (5.8%, 26/447). However, there was no significant difference in the prevalence of *Cryptosporidium* between the cool months (8.9%, 53/594) and the hot months (11.7%, 78/605) (Tellevik *et al.*, 2015).

For *Giardia*, several studies have reported that *G. duodenalis* infections were found to be higher in the wet season than in the dry season (Vargus *et al.*, 2004; Breathnach *et al.*, 2010; Siwila *et al.*, 2011). The results in the current study were opposite to some of the reports above, as I did not find an association of *Giardia* infection with the rainy season. Otherwise, the study in Tanzania found

that *G. duodenalis* infection was more prevalent in the cool months (6.4%, 38/594), than in the hot months (3.0%, 20/665; $p = 0.0048$) but the prevalence of *Giardia* infection was 4.2% (34/812) in the rainy season and 5.4% (24/447) in the dry season, it was not significantly affected by rainfall (Tellevik *et al.*, 2015). The result of this study were different from a study on the characteristics of *G. duodenalis* infection in the Northwest of the United Kingdom which found that *Giardia* assemblage A and B did not significantly change between the four seasons. However, Assemblage B cases tended to be more frequent in winter (dry season) and Assemblage A in summer (wet season) whereas these results found that *Giardia* assemblage A/B was more frequent in August and September (dry season) followed by assemblages B and assemblages A (Minetti *et al.* (2014). However, the difference of the results found maybe explained by the studies sample size being small. In Zambia, a study by Siwila *et al.* (2011) on seasonal variation and incidence of *Cryptosporidium* and *G. duodenalis* in children from four pre-schools in Kafue was carried out Using Direct Immunofluorescence Assay, *Cryptosporidium* spp. was found in 30.7% (241/786) of children whereas *G. duodenalis* was detected in 29.0% (228/786). Both *Cryptosporidium* spp. and *G. duodenalis* infections were higher in the wet season than in the dry season. Cryptosporidiosis infection was symptomatic with patients suffering from diarrhoea whereas patients infected with *Giardia* did not. The majority of studies point to a relationship between seasonality and rainfall and not temperature, so there is some uncertainty as to whether the effects of temperature were considered. The prevalence of *G. duodenalis* was significantly higher in controls than in cases. This is in agreement with other studies in pediatric populations (Kotloff *et al.*, 2013; Mbae *et al.*, 2013, Platts-Mills *et al.*, 2014; Krumkamp *et al.*, 2015).

To summarise, mixed infections with both *Giardia* assemblage A and B (A/B mixed) were predominant in this study. Several previous studies demonstrated that the assemblage B was predominant in some countries, followed by assemblage A and mixed infection, respectively. Although, the ratios of assemblage A, B and mixed infection found in this work disagreed with many published papers, assemblage B was significantly higher than assemblage A. From this study, co-infection between *G. duodenalis* and *Cryptosporidium* spp. was found in many samples. For *Cryptosporidium* spp., the predominance of *C. hominis* from this finding indicated that the anthroponotic route plays a major role in *Cryptosporidium* transmission in Malawi.

CHAPTER THREE

MULTILOCUS GENOTYPING OF *GIARDIA DUODENALIS* AMONG CHILDREN IN CAMBODIA

3.1 INTRODUCTION

3.1.1 Intestinal parasite infections in Cambodia

Gastrointestinal parasites are a major global health problem, especially in developing countries (Lee, 2004). About 1.2 billion people in Southeast Asian (SEA) countries are infected with nematodes and/or soil-transmitted helminths (STH) (WHO, 2008). The prevalence of STH infection is higher than other parasite infections (cestodes and trematodes). Among SEA countries, Cambodia is recognised as one of the poorest with a lack of good health care provision and high rates of malnutrition (Marmot, 2008). Less than a third of people in Cambodia have good quality sanitation facilities (UNICEF, 2009). Due to poor sanitation and socio-economic conditions, children who live in this region have a high risk of infection from intestinal parasites. While mortality caused by parasitic infections is low in children (ages ranging from 0-16 years), STH significantly impact on their growth, nutrition and cognitive development (Brooker *et al.*, 2008; Hotez *et al.*, 2008).

There are published data outlining the intestinal parasitic infections of children and refugees in Cambodia (Lurio *et al.*, 1991; Gyorkos *et al.*, 1992; Lee *et al.*, 2002; Moore *et al.* 2012; Moore *et al.*, 2015). There are limited facilities for detection of parasitic infection in Cambodia outside of the capital city Phnom Penh. The Angkor Hospital for Children (AHC) in Siem Reap Province is a NGO-run hospital which sees approximately 400 outpatients per day. The hospital laboratory also serves the satellite paediatric ward at Sotr Nikom District Hospital (SDH), about 40 kilometres from Siem Reap town. Direct faecal microscopy for parasites and culture for *Salmonella* species and *Shigella* species is routinely performed at the hospital. Between January 2006 and September 2011, parasites were detected in 3,121 (19.1%) of 16,372 faecal samples from children examined at the AHC. Common parasites were *Giardia duodenalis* (8%), hookworm (5.1%) and *Strongyloides stercoralis* (2.6%). *Giardia* and hookworm infections increased significantly over the five years. *Cryptosporidium* was rarely investigated (Moore *et al.*, 2012). There is a lack of knowledge of the

true clinical relevance of infection with *Strongyloides* and *Cryptosporidium* and the different infection patterns associated with disease. Recently, Moore *et al.* (2015) investigated gastrointestinal parasite infections in children with diarrhea who attended AHC. Stool samples were collected in 2012 and were examined by using the concentration technique and the microscopic method. The agar plate culture and the charcoal copro-culture method were used for hookworm and *S. stercoralis*. The overall prevalence of parasitic infection was 39.3%. The highest prevalence of parasitic infection was hookworm (14.4%), followed by *S. stercoralis* (11.6%) and *G. duodenalis* (11.2%). Children within the 1-5 years age group were most commonly infected with *G. duodenalis* whilst children at the age of more than 5 years were commonly infected with hookworm and *S. stercoralis*.

3.1.2 Molecular genotyping for *Giardia* in Cambodia and countries in Southeast Asia

It is important to understand the taxonomy and epidemiology for *Giardia* sub-assemblages. This is because an understanding of the genotypes/assemblages could help in the treatment and prevention of further cases of *Giardia*. There are a number of different *Giardia* species and genotypes, by understanding the epidemiology of *Giardia* it is possible to study whether cross species transmission can occur as well as the risk that the species particular assemblages offer from exposure. It can also help understanding of the potential for zoonotic infections from animal to human. Molecular typing tools enable a full picture of *Giardia* epidemiology. The molecular characterisation of *Giardia* in humans has been described using a PCR-based typing technique. Previously, Small subunit ribosomal ribonucleic acid (SSU-rRNA) was used to classify *Giardia* from other parasite species and also used for characterisation of *G. duodenalis* assemblages. The SSU-18S rRNA gene is one of the most frequently used genes in phylogenetic studies and an important marker for random target PCR in environmental biodiversity screening. However, using only one locus or one marker may not be completely successful in all specimens. Therefore, the use of multilocus genotyping (MLG) has become a popular tool for genotyping and sub-genotyping of this parasite (Amar *et al.*, 2002; Breathnach *et al.*, 2010).

Little is known about the distribution of *Giardia* in humans and animals in Cambodia, especially regarding genotypes (assemblages and sub-assemblages) and also multi-locus genotypes.

However, there were studies of the genetic diversity in some countries in Southeast Asia such as analysis of PCR-RFLP at *tpi* gene in Indian patients; the prevalence of *Giardia* infection both in giardial diarrhoea (80%, 362/452) and asymptomatic giardiasis (94%, 425/452) was highly associated with Assemblage B. There were correlations between diarrhoea and assemblage A subgroup II alone or dual infections with both assemblage A and B (Ajjampur *et al.*, 2009). Genotyping of *G. duodenalis* was undertaken in 61 faecal specimens (with and without diarrhoea) from Bangkok and in rural areas of Thailand; assemblage B was found in 51%, mixed infection of both assemblages was detected in 41% and the assemblage A alone was detected in 8% by using the *tpi* gene. The prevalence of assemblage B and mixed infection (A/B mixed) were significantly higher than assemblage A alone. The sub-assemblage AI was found in 12%, the sub-assemblage AII was found in 88% by using the *bg* gene. Using the *gdh* gene, the sub-assemblage BIII was detected in 45%, and the sub-assemblage BIV was identified in 54.5%. Interestingly, the sub-assemblage AI was found in 100% from symptomatic cases and the sub-assemblage BIII sub-assemblages was detected in 50% (Tungtrongchitr *et al.*, 2010). A previous study performed by Huey *et al.* (2013) based on multilocus analysis revealed that 42% of the Orang Asli isolates belong to assemblage A and 45% belonged to assemblage B.

3.1.3 Using a multi-locus sequence typing (MLST) for *Giardia duodenalis* infection in Cambodia and countries in Southeast Asia

G. duodenalis is a highly complex species and it has been categorised into at least eight (A to H) morphologically identical genotypic assemblages with variable host occurrence. Humans are infected by assemblages A and B that demonstrate extensive genetic variability at the sub-assemblage level. MLST of structural and housekeeping genes is a useful method to detect potentially zoonotic genotypes of this pathogen. Genotyping of human and animal isolates from the same area, in conjunction with epidemiological data, can provide crucial information about transmission routes and epidemiological differences between *G. duodenalis* assemblages.

However, there is a lack of information on the distribution and diversity of multi-locus genotypes of *Giardia* in humans from developing countries, since only a few studies implemented a multi-locus genotyping approach based on the sequencing of the *bg*, *gdh* and *tpi* loci (Ajjampur *et al.*, 2009; Tungtrongchitr *et al.*, 2010). The number of studies relating to *Giardia* in Cambodia is also small which limits any firm conclusions that can be drawn. Much of the available literature is confined only to general information on intestinal parasites such as the prevalence of infections with *Giardia* (Moore *et al.*, 2012; Moore *et al.*, 2015). Moreover, several other studies reviewed were based upon retrospective studies and microscopic methods (Mak, 2004; Smith and Caccio, 2006; Moore *et al.*, 2012). As such information on *Giardia* infection in Cambodia was limited, primarily due to the lack of quality instruments and available methods for detection (Dorny *et al.*, 2011; Jeon *et al.*, 2011; Khieu *et al.*, 2014a; Khieu *et al.*, 2014b; Miyamoto *et al.*, 2014). Similarly, there were few reports available that investigated the risk factors of *G. duodenalis* assemblages and/or the association between *Giardia* assemblages and the clinical symptoms of giardiasis.

Previous studies have described the clinical outcomes of this parasite and its assemblages (Abdel-Moniem and Sultan, 2008; Sahagun *et al.*, 2008; Mohammed Mahdy *et al.*, 2009). However, there is limited information about *Giardia* infection in several Asian countries. Moreover, no study in Cambodia has used an MLST technique to determine the molecular characterisation of *Giardia* at multiple loci in a larger group of hosts. Therefore, we identified the assemblages and sub-assemblages of *G. duodenalis* from stool samples of children in various regions of Cambodia admitted at Angkor Hospital for Children, Sotr Nikom District Hospital and The Lake Clinic.

3.2 AIMS OF THE STUDY

The aim of this study was to investigate the distribution of *G. duodenalis* genotypes (assemblage and sub-assemblage prevalence and zoonotic genotypes) in faecal specimens obtained from children under 16 years of age living in different locations in Cambodia. A standard MLST method based on the sequencing of the *tpi*, *bg* and *gdh* loci was used to characterise the *G. duodenalis* genotypes.

The three hospitals which were selected comprise the AHC, SDH and The Lake Clinic (TLC). The three hospitals were selected so that they were representative of the northern (Oddar Meanchey),

northwestern (Banteay Meanchey and Siem Reap), western (Battambang) central (Kampong Thom) and southeastern (Kampong Cham) of Cambodia respectively. Moreover, the three regions represent different sizes and the types of hospital, the types of patients (in-patients and out-patients), and patients came from various locations (both rural and urban areas were included). A large sample size was able to be collected from the representative regions.

The hypotheses of our study are that: 1) there will be anthroponotic species in urban areas 2) there will be zoonotic species in rural areas 3) there will be mixed infection (anthroponotic and zoonotic species) in urban and rural areas. Appropriate markers and/or multilocus genotyping were used to identify assemblages and sub-assemblages of *G. duodenalis*. Our study also assessed the extent of variation in sequences from the isolates within each assemblage. Parasite isolates were typed by using multilocus genotyping, Nested-PCR and/or sequencing of Triose phosphate isomerase gene (*tpi* gene), *bg*, and glutamate dehydrogenase genes (*gdh* gene) for *Giardia*.

The questions addressed in this study are:

1. Are the distributions of the *Giardia* assemblages/sub-assemblage infections different when using *tpi*, *bg* and *gdh* genes? (see topic 3.4.1 & 3.4.2)
2. How does the patients gender, age group, type of diarrhoea, abdominal pain, background of patients and nutritional status affect the *Giardia* assemblages/sub-assemblages? (see topic 3.4.3)
3. What effect do the different *Giardia* assemblages/sub-assemblages have on the duration of the diarrhoea? (see topic 3.4.3 & 3.4.4)
4. How do the different *Giardia* assemblages/sub-assemblages effect the duration of abdominal pain? (see topic 3.4.3 & 3.4.5)
5. Are there different distributions of *Giardia* assemblages/sub-assemblages and infections in different geographical locations (areas)? (see 3.4.6)

3.3 MATERIALS AND METHODS

3.3.1 Source of samples

This research was in collaboration with Catrin Moore¹, Christopher Parry², Nicholas Beeching³, Wendi Bailey⁴, Varun Kuma⁵, Caroline Corless⁶, and Rachel Chalmers⁷. It was some of the members of this group who collected the faecal specimens. Briefly, since 1st March/April 2012 to the 31st August 2012, stool samples were collected from children less than 16 years of age with acute diarrhoea who presented themselves to the Cambodian AHC, SDH and TLC from March/April 2012-August 2012.

The Cambodian Angkor Hospital provides free health care to children up to and including the age of 16 years. The hospital is funded by charitable donations and its 50 beds mainly admit patients from Siem Reap and surrounding provinces, although some patients do travel from all over Cambodia. Annually the hospital sees about 125,000 visits and admits about 4000 patients. The hospital collects approximately 5,000 stool samples every year and these are screened by direct microscopy for intestinal parasites (Moore *et al.*, 2012). The samples used in this study have been previously screened for *Giardia* by microscopy (Moore *et al.*, 2015). However, they have not been examined for *Giardia* by using multilocus analysis. Therefore, the samples were screened for the presence of this parasite by the PCR based method, and multilocus genotyping was used to identify assemblages and sub-assemblages of *G. duodenalis*.

¹ = The Pneumococcal Surveillance Group, John Radcliffe Hospital, Oxford

² = The National Centre for Zoonosis Research, Leahurst, Chester High Road

³ = Tropical and Infectious Disease Unit, Royal Liverpool University Hospital

⁴ = Diagnostic Laboratory, Liverpool School of Tropical Medicine, University of Liverpool

⁵ = Angkor Hospital for Children, Siem Reap, Cambodia

⁶ = Infection and Immunity, Royal Liverpool and Broadgreen University Hospitals NHS Trust, Liverpool,

⁷ = *Cryptosporidium* Reference Unit, Public Health Wales Microbiology ABM, Singleton Hospital, UK.



Figure 3.1 Map of Cambodia showing the regions of Oddar Meanchey (1), Banteay Meanchey (2), Battambang (3) Siem Reap (4), Kampong Thom (5) and Kampong Cham (6). Adapted from: <https://heifer12x12.com/map-of-cambodia/>

3.3.2 DNA samples

Five hundred and six DNA samples were obtained from the Pneumococcal Surveillance Group, Microbiology Level 7, John Radcliffe Hospital, Oxford. The DNA had been extracted, using the QIAamp DNA Stool mini kit (Qiagen), as part of previous studies of faecal parasitic infections. DNA was stored at -80°C until shipped on dry ice to the University of Liverpool where it was stored at -20°C prior to analysis. They included all the samples known to be direct microscopy-positive for *Giardia* and samples which were direct microscopy negative. The samples were chosen randomly without knowledge of the screening results for *Giardia* by microscopy.

The microtubes contained DNA samples which were checked to see if they had enough volume to carry out PCR. DNA samples which had a volume of at least 10 µl were chosen to be studied as we had to make sure that sufficient volume was available so that we could carry out PCR for all

experiments (SSU-rRNA, *tpi*, *bg*, and *gdh* gene). Tubes without sufficient volume were left in the box and some which had been fully/almost fully utilized were disposed off. This left a total of 484 DNA samples for this study.

3.3.3 Clinical information, for examples gender, date of collection and date of birth (see Appendix) was recorded by the researchers who collected the faecal specimens from AHC, SDH and TLC in Cambodia. Samples from children who attended the out-patient clinic of the AHC, between the 3rd April and the 29th June 2012 were collected. The samples were taken when the children presented with diarrhoea, abdominal pain or showed signs of suffering from anaemia or malnutrition. The samples were screened for *Giardia* using microscopy. The duration of diarrhoea, for this study duration is defined as the day the patient attended hospital minus the date the diarrhoea first started. For example, if a patient first attended hospital on the 10th of the month but the diarrhoea had started on the 7th of the month, duration was classed as three days. Most of the samples came from outpatients; few patients were admitted to hospital.

3.3.4 Multilocus genotyping and subtyping of *Giardia*

Primers were selected from published papers for genotyping and subtyping *Giardia* as described in Chapter 2 (2.3.4.1).

3.3.5 Triose phosphate isomerase (*tpi*) gene

PCR was carried out as described in Chapter 2 (section 2.3.4.3).

Identification of *G. duodenalis* assemblages. Three PCRs were performed (*tpi*, *bg* and *gdh* genes). First a screening assay based on the *tpi* locus, specific for assemblages A and B was performed on all samples (484 samples).

3.3.5.1 *Giardia* Assemblage A specific PCR

PCR was carried out as described in Chapter 2 (section 2.3.4.3.1).

3.3.5.2 Assemblage B specific PCR

PCR was carried out as described in Chapter 2 (section 2.3.4.3.2).

3.3.6 *β-giardin* gene

PCR was carried out as described in Chapter 2 (section 2.3.4.4). Samples positive in the *tpi* locus assay underwent a nested-PCR based on the *bg* gene.

3.3.7 Glutamate dehydrogenase (*gdh*) gene

PCR was carried out as described in Chapter 2 (section 2.3.4.5). Samples positive in the *tpi* locus assay underwent a nested-PCR based on the *gdh* gene.

3.3.8 Purification of PCR products

All PCR products were purified as described in Chapter 2 (2.3.6).

3.3.9 Sequencing of PCR products and DNA Sequence analysis

The purified PCR products were sequenced using the methodology previously described in Chapter 2 (2.3.7).

3.3.10 Statistical analyses

Associations between *Giardia* infection and the child's gender, age and other clinical information were analysed. In order to analyse the correlation among those as mentioned above, Chi-square tests were performed using SPSS (version 7.0). The age data were divided into three groups (less than 5 year, 5-10 years, and 10-16 years). Comparison of proportions of assemblage A and B were tested by Fisher's exact test. The level of statistical significance was set at $p < 0.05$.

3.3.11 Ethical consideration for the study

Ethical approval for this study was approved by the Institutional Review Board at AHC, the Oxford Tropical Research Ethics Committee (OXTREC12-12) and the Public Health Wales Research Risk Review Committee.

3.4 RESULTS

3.4.1 Overall prevalence of *G. duodenalis* assemblages

The *G. duodenalis* assemblage was identified in 74 of the 484 faecal specimens (15.4%) either by sequencing of the *tpi* gene or by on-gel visualisation of the *tpi* assemblage-specific products. The most common *Giardia* assemblage was Assemblage B, 83.8% (62/74); Assemblage A was found in 10.8% (8/74) and 5.4% (4/74) demonstrated a mixed assemblage (determined as a sample demonstrating the sequences of two different assemblages at different loci and/or amplifying the products of both assemblages following the *tpi* assemblage-specific PCRs). Using *bg* locus, Assemblage B was found in the majority of the specimens (79.2%, 38/48), followed by Assemblage A which was found 8.3%, 4/48) and 6.3% (3/48) showed a mixed assemblage infection (A/B mixed). For *gdh* locus, Assemblage B was found in 86.3% (44/51), followed by Assemblage A which was found in 7.8% (4/51) and 5.9% (3/51) showed a mixed assemblage A and B infection.

Of the four specimens showing a mixed infection, one was diagnosed by the *tpi* assemblage A and B-specific PCR while it was typed as B at both the *bg* and *gdh* loci, one was diagnosed by the *tpi* assemblage A and B-specific PCR whereas it was typed assemblage A at both the *bg* and *gdh* loci, one was typed assemblage A and B-specific at the *tpi* locus (while assemblage B at the *bg* locus and data not available at the *gdh* locus (did not succeed for sequencing), and one was typed as A and B-specific at the *tpi* locus whereas it did not succeed for sequencing at both the *bg* locus and the *gdh* locus.

Table 3.1 The prevalence of *G. duodenalis* assemblages examined in 484 faecal samples by nested-PCR

Markers	Sample tested	Positive samples by PCR	Sequenced	Ass. A	Ass. B	Ass. AB	Not determined
<i>Tpi</i>	484	15.5% (74/484)	74	10.8% 8/74)	83.8% (62/74)	5.4% (4/74)	0
<i>Bg</i>	74	64.9% (48/74)	48	8.3% (4/48)	79.2% (38/48)	6.3% (3/48)	3
<i>Gdh</i>	74	68.9% (51/74)	51	7.8% (4/51)	86.3% (44/51)	5.9% (3/51)	0

Table 3.2 Assemblage typing results by three different markers in the four samples with a mixed assemblage A and B infection

ID of samples	<i>tpi</i>	<i>Bg</i>	<i>gdh</i>
15	A+B	B	B
79	A+B	A	A
411	A+B	B	n/a
1074	A+B	n/a	n/a

3.4.2 *G. duodenalis* sub-assemblages and genotypes by locus

3.4.2.1 Triose phosphate isomerase (*tpi* gene)

The 332 bp fragment of the *tpi* gene (Assemblage A) and the 400 bp fragment of the *tpi* gene (Assemblage B) of 484 DNA samples were amplified. The overall prevalence of *G. duodenalis* by the *tpi* gene was 15.5% (Table 3.1). Out of 74 samples, 74 (100%) were successfully sequenced and genotyped. The results of the sequencing analysis matched the GenBank reference (Table 3.1). The infection rates of assemblage B and assemblage A were significantly higher than mixed infection of assemblage A and B. Of the 74 samples, 62 (83.8%, 62/74) were infected with assemblage B, eight (10.8%, 8/74) with assemblage A and four (5.4%, 4/74) with assemblage A/B (Table 3.1). By using the *tpi* gene, almost all assemblage A belonged to the sub-assemblage AII (92%, 11/12). Whilst Assemblage B parasites, 43% (12/28) belonged to sub-assemblages BIII, 43% (12/28) displayed assemblage VB906855 and 11% (3/28) belonged to B heterogenous (Figure 3.2).

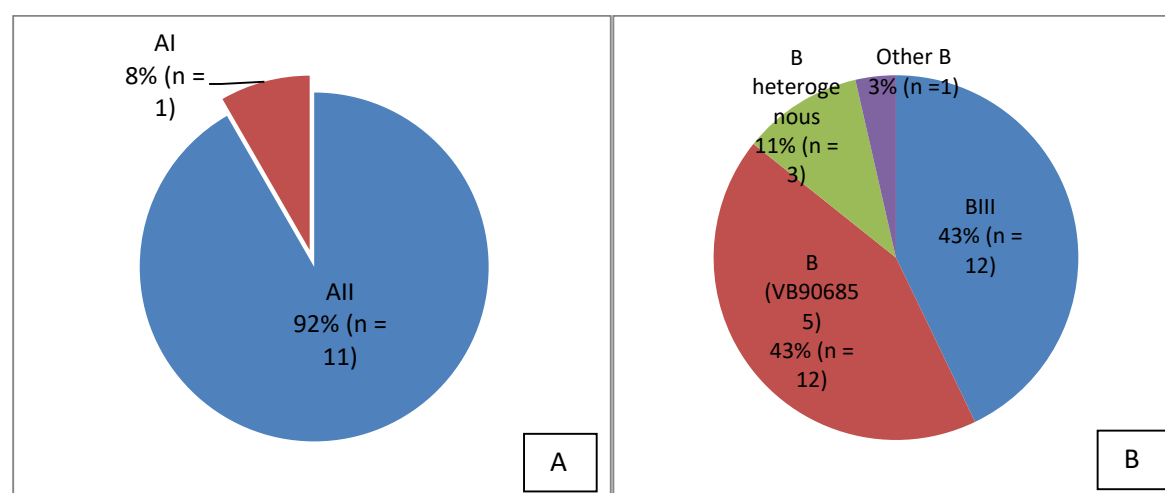


Figure 3.2 Overall sub-assemblage typing results at the *tpi* locus in 12 assemblage A (A) and 28 assemblage B isolates (B). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity.

3.4.2.2 *Beta-giardin* gene

The 511 bp fragment of the *β-giardin* gene of 74 DNA samples were amplified. The overall prevalence of *G. duodenalis* by the *bg* gene was 64.9% (48/74) (Table 3.1).

Out of 48 positive samples, 48 (100%) were successfully sequenced and genotyped. The results of the sequencing analysis matched the GenBank reference (Table 3.1). Of the 48 samples, 38 (79.2%, 38/48) were infected with assemblage B, four (8.3%, 4/48) with assemblage A and three (6.3%, 3/48) with assemblage A/B (Table 3.1). By using the *bg* gene, 50% (2/4) contained assemblage AI, 50% (2/4) displayed assemblage AII. For Assemblage B, 40% (6/15) contained assemblage BIII, 40% (6/15) showed assemblage other B and 20% (3/15) belonged to B heterogenous (Figure 3.3).

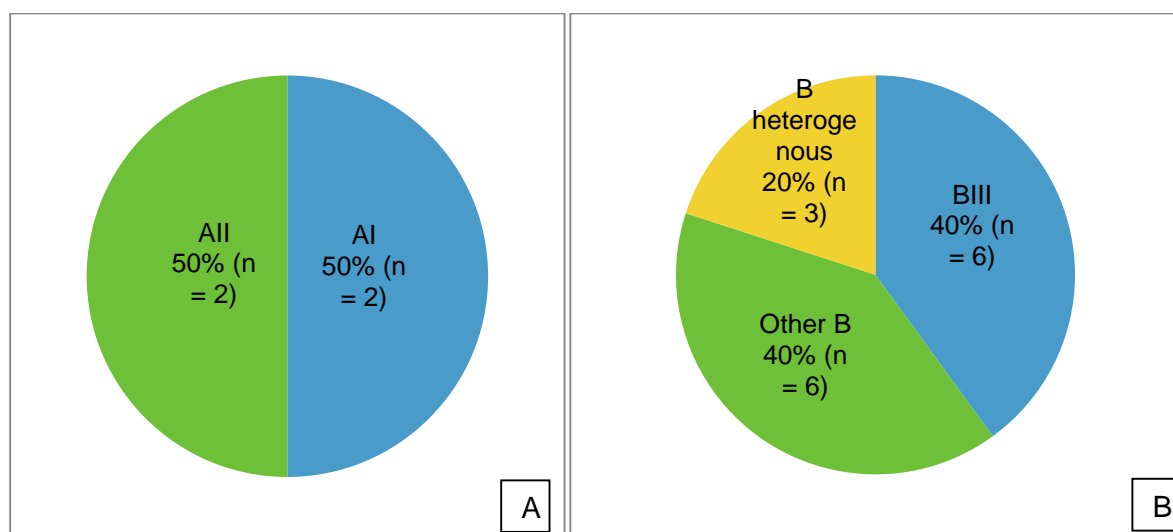


Figure 3.3 Overall sub-assemblage typing results at the *bg* locus in 4 assemblage A (A) and 15 assemblage B isolates (B). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity.

3.4.2.3 *Glutamate dehydrogenase* gene

The 530 bp fragment of the *gdh* gene of 74 DNA samples was amplified. The overall prevalence of *G. duodenalis* by the *gdh* gene was 68.9% (51/74) (Table 3.1). Of the 51 samples, 44 (86.3%, 44/51) were infected with assemblage B and four (7.8%, 4/51) with assemblage A and three (5.9%, 3/51) with assemblage A/B (Table 3.1). By using the *gdh* gene found 50% (19/28) contained sub-assemblage BIV H43 (EF507682), 24% (9/28) displayed sub-assemblage BIV (KC96064.1), 16% (6/28) belonged to sub-assemblage B (VAN/90/UBC/54) and 16% (6/28) belonged to sub-assemblage other B (Figure 3.4).

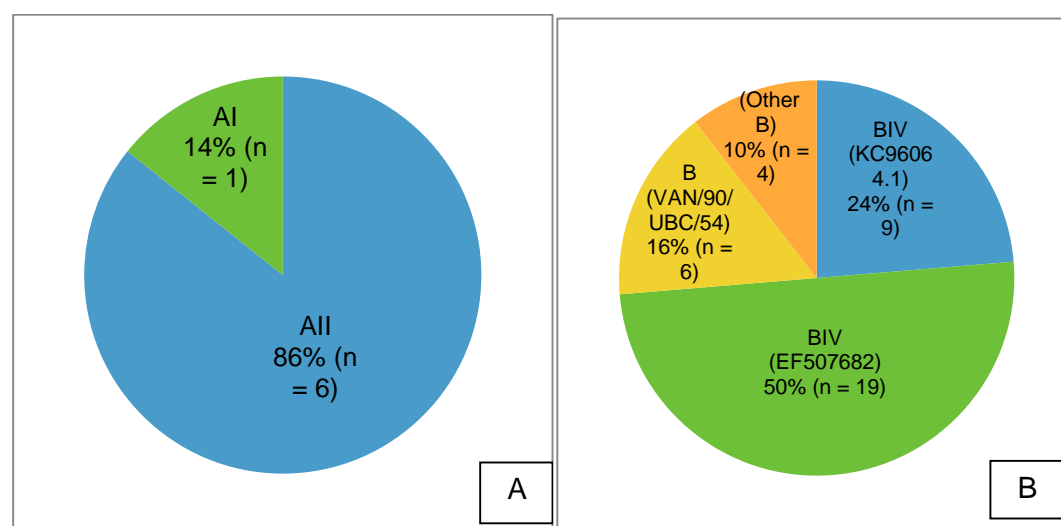


Figure 3.4 Overall sub-assemblage typing results at the *gdh* locus in 7 assemblage A (A) and 38 assemblage B isolates (B). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity.

3.4.3 Associations between patient characteristics and *G. duodenalis* infection in children

In this study, of 74 *G. duodenalis* positive samples, 16.2% were from males and 14.3% were from females. There was no statistical difference of the proportion of *G. duodenalis* infections between male and female. Regarding age, children infected with *G. duodenalis* were most commonly

identified in the age >5 to 10 years (18.5%), followed by the age group 10-15 years (14.3%) and the age group <1 to 5 years 12.8%), respectively. However, these differences were not statistically significant (Table 3.3 and Table 3.4).

Table 3.3 demonstrates the characteristics (gender and age group) of children infected with *Giardia* by assemblages A, B and AB, although the sample size was small, assemblage B was the most common of the three assemblages, both in the male samples and the female samples. Assemblage A and assemblage AB were only found in patients under the age of 10, whereas patients aged over 10 were only found to be infected by assemblage B. The patients between the ages of 5-10 years had the highest prevalence of *G. duodenalis* (18.5%) infection whereas the patients aged between 10 to less than 16 years and the aged less than 5 years (1-5 years) were 14.3% and 12.8%, respectively. Assemblage B infection was especially found in the children at the age from 10 to less than 16 years (100%) the most, followed by the age groups between 5 -10 years (85.3%), and at the age from 1-5 years (72%), respectively.

Table 3.3 Characteristic of children infected with *Giardia (tpi gene)* (n=74)

Patient characteristics	Prevalence of <i>G. duodenalis</i> infection (%)	Ass. A	Ass. B	Ass. AB
Gender				
Male	42/260 (16.2%)	3 (7.1%)	36 (85.7%)	3 (7.1%)
Female	32/224 (14.3%)	5 (15.6%)	26 (81.3%)	1 (3.1%)
Age (years)				
less than 5 years	25/195 (12.8%)	5 (20%)	18 (72%)	2 (8%)
5-10 years	34/184 (18.5%)	3 (8.8%)	29 (85.3%)	2 (5.9%)
10-16 years	15/105 (14.3%)	0 (0%)	15 (100%)	0 (0%)
Total	74/484	8	62	4

Table 3.4 shows that sub-assemblage BIV H43 (EF507682) was much more prevalent in males compared to females, whereas sub-assemblage BIV (KC96064.1) showed broadly similar rates of infection between the sexes. Sub-assemblage BIV H43 (EF507682) showed broadly similar rates of infection between patients of age less than 5 years, 5-10 years and 10-16 years. Infection with *G. duodenalis* resulted in 17% of patients experiencing abdominal pain. Sub-assemblage B (unidentified) was the most common assemblage both for those not complaining pain and those patients that did. Just over two thirds of those infected with *G. duodenalis* did not experience symptoms of diarrhoea. Again sub-assemblage B (unidentified) was the most prevalent assemblage in both those infected that had diarrhoea and those that didn't have diarrhoea. The presence of the parasites was not significantly associated with diarrhoea, abdominal pain status in children but it was significantly associated with being an out-patient (Fisher's exact test, $p = 0.01$) and normal nutrition status (Fisher's exact test, $p = 0.049$). (Table 3.4).

Table 3.4 Associations between patient characteristics and *G. duodenalis*, sub-assemblage infection based on *gdh* gene from children aged

less than age 16 years (n=74)

Patient characteristic	Prevalence of <i>Giardia</i> infection (%)	Ass A	Sub-ass. BIVH43 (EF507682)	Sub-ass. B (VAN/94/UBC/122)	Sub-ass. B (VAN/90/UBC/54)	Sub-ass. BIV (KC96064.1)	Sub-ass. B (un-identified)	Significant(P-value)
Gender: Male Female	42/260 (16.2%) 32/224 (14.3%)	AI =0, AII =4 AI =1, AII =2	11 (26.19%) 3 (9.38%)	2 (4.76%) 3 (9.38%)	3 (7.14%) 3(9.38%)	7 (16.67%) 5 (15.63%)	15 (35.71%) 15 (46.88%)	$P = 0.33$
Age group ; years less than 5 5 - 10 10 – 16	 25/195 (12.8%) 34/184 (18.5%) 15/105 (14.3%)	 AI =1, AII =2 AI =0, AII =4 AI =0, AII =0	 5 (20%) 6 (17.6%) 3 (20%)	 1 (4%) 3 (8.8%) 2 (13.3%)	 1 (4%) 2 (5.9%) 2 (13.3%)	 3 (12%) 6 (17.7%) 3 (20%)	 12 (48%) 13 (38.2%) 5 (33.3%)	 $P = 0.084$ (group 1 & 2) $P = 0.228$ (group 2 & 3) $P = 0.424$ (group 1 & 3)
Out-patients In-patients	55/273 (20.1%) 19/211 (9%)	AI =1, AII =6 AI =0, AII =0	9 (16.4%) 5 (26.3%)	5 (9.1%) 1 (5.3%)	5 (9.1%) 4 (21.1%)	11 (20%) 1 (5.3%)	18 (32.7%) 8 (42.1%)	$P = 0.01$
Abdominal pain Without abdominal pain	58/340 (17%) 16/144 (11.1%)	AI =1, AII =5 AI =0, AII =1	11 (19%) 3 (18.8%)	6 (10.3%) 0 (0%)	4 (6.9%) 2 (12.5%)	9 (15.5%) 3 (18.7%)	22 (37.9%) 7 (43.7%)	$P = 0.061$
Diarrhoea Non-diarrhoea	24/164 (14.6%) 50/320 (15.6%)	AI =1, AII =2 AI =0, AII =4	7 (29.2%) 7 (14%)	3 (12.5%) 3 (6%)	1 (4.2%) 4 (8%)	1 (4.17%) 10 (20%)	9 (37.5%) 22 (44%)	$P = 0.443$
Nutritional status; Malnourished Normal nutrition status	 2/33 (5.3%) 72/446 (16.1%)	 AI =0, AII =0 AI =1, AII =6	 0 14 (19.4%)	 0 6 (8.3%)	 0 6 (8.3%)	 0 12 (16.7%)	 2 (100%) 27 (37.5%)	 $P = 0.049$

3.4.4 The association between the duration of diarrhoea and infection with *G. duodenalis*

The duration of diarrhoea of children infected *G. duodenalis* was generally of a short period with the majority of cases lasting no longer than 10 days and most lasting just one day. Sub-assembly BIV (KC96064) was detected the most at 1 day of diarrhoea duration. The number of each sub-assembly was not significantly different (Table 3.5).

Table 3.5 Frequency of samples infected with *G. duodenalis* (*tpi* gene) on diarrhoea duration days specified by sub-assemblies (AI, AII, BIV H43 (EF507682), B (VAN/94/UBC/122), B (VAN/94/UBC/122), BIV (KC96064.1) and B (un-identified).

Diarrhoea duration days (no. of days)	Frequency of samples infected with <i>G. duodenalis</i>	Ass. A	Sub-ass. BIV H 43 (EF507682)	Sub-ass. B (VAN/94/UBC/122)	Sub-ass. B (VAN/90/UBC/54)	Sub-ass. BIV (KC96064.1)	Sub-ass.B (un-identified)
0	1		1	0	0	0	0
1	12	AII =1	2	3	1	5	0
2	2	AI =1	1	0	0	0	0
3	2	0	1	0	0	0	1
4	2	0	1	0	0	0	1
7	2	AII =1	0	0	0	1	0
10	2	A =1	1	0	0	0	0
30	1	0	1	0	0	0	0
152	1	0	0	0	0	0	1
n/a	49	0	0	0	0	0	0
Total	74	AI =2 AII =2	8	3	1	6	3

3.4.5 The association between the duration of abdominal pain and infection with *G. duodenalis*

Table 3.6 demonstrates that as well as the diarrhoea itself not generally lasting longer than a few days, the accompanying abdominal pain also tends to last for just a few days with the vast majority of cases of pain only lasting between 2-4 days and rarely longer than seven days. For the sub-assemblages recorded, all had the same number at 3 days of abdominal pain duration which were found 2 samples infected with each sub-assemblage (Table 3.6).

Table 3.6 Frequency of samples infected with *G. duodenalis* (*tpi* gene) on abdominal pain duration days specified by sub-assemblages (AI, AII, BIV H43 (EF507682), B (VAN/94/UBC/122), B (VAN/94/UBC/122), BIV (KC96064.1) and B (un-identified).

Abdominal pain duration days (no. of days)	Frequency of samples infected with <i>G. duodenalis</i>	Ass A	Sub-ass. BIV H43 (EF507682)	Sub-ass B (VAN/94/UBC/122)	Sub-ass. B (VAN/90/UBC/54)	Sub-ass. BIV (KC96064.1)	Sub-ass. B (un-identified)
0	1	0	0	0	0	0	1
1	6	AII =1	2	0	1	0	2
2	9	AII =1	2	2	1	1	2
3	12	AII =2	2	2	2	2	2
4	8	0	2	1	2	0	3
5	5	A=1	0	0	0	2	2
6	3	AII =1	1	0	0	0	1
7	4	AI =1	0	0	0	2	1
10	1	0	0	1	0	0	0
14	1	AII =1	0	0	0	0	0
30	3	0	2	0	0	0	1
152	1	0	0	0	0	0	1
304	1	0	0	0	0	0	1
730	1	0	0	0	0	0	1
Total	56	A =1, AI =1, AII =6	11	6	6	7	18

3.4.6 The association between the different geographical locations and infection with *G. duodenalis* (*tpi* gene)

Table 3.7 demonstrates the distribution of assemblages of *Giardia* from each area in Cambodia. The overall prevalence of *G. duodenalis* was 15.3%. The prevalence of *G. duodenalis* recorded from

samples collected in Kampong Thom was the most predominant (19.2%, 5/26), followed by Battambang (17.4%, 4/23), Siem Reap (15.8%, 53/336, Kampong Cham (14.3%, 2/14), Banteay Meanchey (12.9%, 8/62), and Oddar Meanchey (8.7%2/23), respectively. There were no significant differences between the prevalence of *G. duodenalis* by *tpi* among those from each region in Cambodia (Fisher 's exact test, $p > 0.05$) (Table 3.7).

With reference to *G. duodenalis* infection, the assemblage B, assemblage A and AB mixed can be found only in Siem Reap which were found to be 83 % (44/53), 9.4% (5/53) and 7.5% (4/53), respectively. The assemblage B was the most prevalence in Siem Reap. The infection rates of assemblage B were significantly higher than assemblage A. and assemblage AB (Fisher 's exact test, $p = 0.013$). While Battambang found the assemblage A and assemblage B at the same rates (50%, 2/2). Kampong Thom found the assemblage B (80%, 4/5) more than assemblage A (20%, 1/5). No mixed infections were identified in Battambang and Kampong Thom. For the rest of the regions were assemblage B was only found as follows, Banteay Meanchey (100%, 8/8), Kampong Cham (100%, 2/2) and Oddar Meanchey (100%, 2/2)(Table 3.7).

Table 3.7 The association between the different geographical locations and infection with *G. duodenalis* (*tpi* gene) by Assemblages A, B and AB mixed (n = 74)

Regions (Provinces)	Prevalence of <i>G. duodenalis</i> infection (%)	Ass. A	Ass. B	Ass. AB
Siem Reap	53/336 (15.8%)	5 (9.4%)	44 (83%)	4 (7.5%)
Battambang	4/23 (17.4%)	2 (50%)	2 (50%)	-
Kampong Thom	5/26 (19.2%)	1 (20%)	4 (80%)	-
Banteay Meanchey	8/62 (12.9%)	-	8 (100%)	-
Kampong Cham	2/14 (14.3%)	-	2 (100%)	-
Oddar Meanchey	2/23 (8.7%)	-	2 (100%)	-
Total	74/484 (15.3%)	8/74 (10.8%)	62/74 (83.8%)	4/74 (5.4%)

3.5 DISCUSSION

The study showed that the method with the highest sensitivity for diagnosing *G. duodenalis* was *tpi* (15.5%, 74/484). Of the 74 human specimens successfully sequenced and genotyped using the *tpi* gene, only 51 (68.9%, 51/74) and 48 (64.9%, 48/74) samples were successfully amplified by *gdh* and *bg* loci, respectively (74 = positive samples tested by *tpi*, 100%). This finding is similar to a study by Bertrand *et al.* (2005) that reported the *tpi* PCR was more able to discriminate assemblage and it has higher sensitivity than the *gdh* PCR. In addition, this result corresponds to a study in India by using multilocus genotyping (3 different markers; *tpi*, *bg*, and *gdh*). The result found that *tpi* PCR-RFLP could detect *G. duodenalis* from all positive samples (74/74, 100%), whilst the samples were detected in 62.2% (46/74) and 56.8% (42/74) of cases by using *gdh* and *bg*, respectively. This result disagreed with Geurden *et al.* (2009a), which found a successful amplification of 93% (67/72) using the *tpi* assemblage A and B specific PCR assay from human samples whilst all the samples were amplified by the *bg* locus. Moreover, the differential amplification success of these loci was reported by Broglia *et al.* (2013). The differential amplification rates were 92.4% at the *bg* gene, 44% at the *tpi* gene and only 25.7% at the *gdh* gene. It is unlikely that there will be little variance between the gene copy numbers *tpi*, *bg* and *gdh* markers as they are all single copy genes. Nucleotide mismatches between different PCR primers and the genomic sequences have been suggested for the reason that certain loci lack a clear signal (Broglia *et al.*, 2013). Where mismatches occur this may result in particular primers being unable to amplify distinct loci, for example *bg* primer when presented with some assemblage B sub-types (Robertson *et al.*, 2007a).

From these results, the assemblage characterisation of the cysts from 74 positive samples showed that 83.8% (62/74) of the cysts were assemblage B, 10.8% (8/74) were assemblage A and 5.4% (4/74) were a combination of assemblage A and B. This finding corresponds to a study in Thailand in which the infection rate of assemblage B was higher than assemblage A (Tungtrongchitr *et al.*, 2010). Moreover, these results are consistent to studies from various regions around the world both in developing and developed countries such as Bangladesh (Haque *et al.*, 2003), India (Ajjampur *et al.*, 2009), The Philippines (Yason and Rivera, 2007), England (Amar *et al.*, 2003), The Netherlands (Van der Giessen *et al.*, 2006) and Brazil (Kohli *et al.*, 2008). In contrast, the results were not similar to some published papers from Korea (Yong *et al.*, 2002) and Mexico (Ponce-Macotella *et al.*, 2002).

which found that assemblage A was higher than assemblage B. However, in the current study it can be concluded that assemblage A and assemblage B are the most common in humans in this setting in Cambodia.

The current study demonstrated that most *G. duodenalis* infections were identified in patients between the ages of 5-10 years (18.5%, 34/184) whereas the age from 10 to less than 16 years and the patients aged between 1-5 years and at were 14.3% (15/105) and 12.8% (25/195), respectively. Assemblage B infection was especially found in children from the age from 10 to less than 16 years (100%, 15/15), followed by those of aged between 5-10 years (85.3%, 29/34), and at the aged between 1-5 years (72%, 18/25), respectively. These results are similar to a study in Malaysia that found the infection of assemblage B at the higher rates in children less than 12 year of age (Mohammed Mahdy *et al.*, 2009).

This study was different from other studies that have taken place in industrialised countries such as Germany. These studies have shown Giardiasis rates of infection to first peak at around the age of two years, the pattern then shows a significant lessening of the rate until a second, stronger peak occurs in young people of 20-30 years of age. (Robert Koch-Institut, 2013). The age distribution of *G. duodenalis* infection is different in endemic regions: in Nicaragua, Guatemala and Cote d'Ivoire, *G. duodenalis* prevalence showed a continuous fall from young children to adults (Téllez *et al.*, 1997; Cook *et al.*, 2009; Schmidlin *et al.*, 2013). However, after a study of children from Uganda and a survey conducted in Niger, *G. duodenalis* infection rates reached a peak in children aged between 3-6 years, which also correlates to studies that have taken place in Rwanda. Peak prevalence in late pre-school age or early school age has also been observed in Nigeria, Zimbabwe, and Lesotho (Oyerinde *et al.*, 1977; Mason and Patterson, 1987; Esrey *et al.*, 1989). The causation of these two patterns in endemic regions, i.e., gradual decline as subjects get older or a peak in infection rates at the age of at 3-6 years and then a subsequent decline, is described and discussed in Chapter 2 (Heimer *et al.*, 2015).

About the role of seasonality, I did not investigate this topic for this chapter as the samples were collected only 6 months (since 1st March/April 2012 to the 31st August 2012). However, some reported indicated that infection rates of *Giardia* infection did not differ with any significance

between the dry season and the rainy season. Cysts can tolerate extremes of temperature so infections can occur in both seasons (Hussein, 2010; Gatei *et al.*, 2006; Siwila *et al.*, 2011). Increasing levels of hygiene and sanitation, together with an understanding of local practices and cultures and education is the best way forward to limiting outbreaks of Giardiasis infections. Understanding the risk factors can inform control measures (Anim-Baidoo *et al.*, 2016).

This study revealed a high association between symptomatic diarrhoea and assemblage B. This finding is consistent with previous studies carried out in human samples from the Netherlands, where *Giardia* assemblage B was detected in 65% of infected samples and assemblage A in 35% of infected samples (van der Giessen *et al.*, 2006). This result corresponded to the study in India, where the greatest prevalence of the *G. duodenalis* assemblage was assemblage B in both adults and children (82.4%), assemblage AII was found in only 9.4% of patients whilst assemblage AI remained undetected (Laishram *et al.*, 2009).

In contrast, previous studies reported a significant association between assemblage A and the presence of symptoms and assemblage B with asymptomatic giardiasis. For example, assemblage A was found from 29 symptomatic cases (100%, 29/29) whilst there was no assemblage B (0/29) in any cases of a study in Bangladesh (Haque *et al.*, 2005). In another study from Peru, assemblage A was found from 19 cases who had diarrhoea whereas assemblage B were found from 6 humans without diarrhoea (Perez Cordon *et al.*, 2008). Elsewhere, a study in Thailand found that all symptomatic cases infected assemblage AI and 50% of patients infected assemblage BIII. The present study has also showed no significant difference between assemblage B and clinical symptoms according to age groups. This study has also demonstrated that assemblage B is highly associated with symptomatic giardiasis. Moreover, assemblage B and AB were more prevalent in males than females. However, this can depend on the status of the patient, the level of host susceptibility, dietary status and the immunity of hosts, as opposed to the particular *Giardia* assemblage (Sahagun *et al.*, 2008).

Interestingly, among nine isolates from the present study, four isolates were identified as other B by the *bg* gene whereas two samples were classified as BIV H43 (EF507682) and the rest of them were detected as BIV (VAN/90/UBC/54) and BIV (KC96064.1) by the *gdh* gene. Two samples were

classified as BIII by *bg* while one sample was identified as BIV (KC96064.1) and another was BIV (VAN/94/UBC/122). The study showed that the method with the highest sensitivity for diagnosing *G. duodenalis* was *tpi* (15.5%, 74/484).

Three samples were characterised as B (heterogenous) by *bg*, whilst two samples were identified as BIV (VAN/94/UBC/122) and another was classified as BIV (KC96064.1) by *gdh*.

Sub-assemblage BIV (VAN/90/UBC/54) was seen more frequency in children who had abdominal pain and less than 10 years of age.

Sub-assemblage BIV H43 (EF507682) was found the most in children who had diarrhoea and less than 5 year ages, followed by the age between 5-10 years and the age between 10-15 years, respectively. Moreover, this sub-assemblage was seen to show a high frequency in children who had abdominal pain (1-6 days) and children who had both abdominal pain around the umbilical area and more generalized pain.

Sub-assemblage BIV (KC96064.1) was detected in children without diarrhoea (11/12) but they had abdominal pain (9/12) around umbilical area the most (5/12) and generalized (1/12). and this sub-assemblage was seen in children at the age of between 5-10 years (6/12), followed by the children at the age of less than 5 years (3/12) and at the age of between 10-15 years (3/12).

Sub-assemblage BIV (VAN/94/UBC/122) was found in 50% of children who had diarrhoea (3/6) and without diarrhoea (3/6). However, all samples who infected with this sub-assemblage had abdominal pain (6/6) which were found around umbilical area the most (4/6), right upper quadrant (1/6) and generalized (1/6).

My hypotheses regarding genetic variation of *Giardia* from the different regions in Cambodia were that: 1) there will be anthroponotic species present in urban areas; 2) there will be zoonotic species present in rural areas; and 3) there will be mixed infection (anthroponotic and zoonotic species) in both urban and rural areas. Characterisation of the *Giardia* assemblage demonstrated that 83.8% (62/74) of the cysts were assemblage B, whereas Assemblage A was identified in 10.8% (8/74) and the mixed assemblage A and B was detected in 5.4% (4/74). The predominance of Assemblage B was found in Siem Reap town (urbanized) and other districts of Siem Reap (rural areas). Moreover, the Assemblage B was found in some other provinces such as Oddar Meanchey, Banteay Meanchey, Battambang, Kampong Thom and Kampong Cham that were representative of the northern,

northwestern, western, central and southeastern of Cambodia, respectively. These data indicate that the anthroponotic route plays a major role in transmission in both urban and rural areas in Cambodia. Furthermore, mixed infections (anthroponotic and zoonotic species) were found in both urban and rural areas; assemblages A and B were found in 4 samples (5.4%, 4/74), including 2 samples were from Siem Reap town and the rest were from rural areas. Thus both anthroponotic and zoonotic species were found in these regions, although the sample size was small. The predominance of anthroponotic transmission has implications for the control of *Giardia* infection in Cambodia. Improvements in hygiene and sanitation should be prioritised, rather than, for example, changes in animal meat processing.

In a study area of rural southern Rwanda, there was not significant evidence that a low socio-economic status associates with a higher prevalence of infection. In rural (not) urban school children, infection rates declined significantly with age. *G. duodenalis* prevalence among children declined throughout school-age. The data suggests that *G. duodenalis* infection is believed to be a common cause of stunting in schoolchildren (Heimer *et al.*, 2015).

Many previous studies found differences in the distribution of *G. duodenalis* between each geographical area. Even though this study found a high association between assemblage B and clinical manifestation of giardiasis, it is not yet possible to conclude that assemblage A is correlated more with asymptomatic giardiasis in Cambodia. However, it demonstrated that assemblage A and B are also common in Cambodia. It is not possible to say with any certainty that the different assemblages in Cambodia exhibit more virulence over one than another. Many local communities have high *Giardia* infection rates, perhaps explained by poor sanitation, including the consumption of fresh fruit that has been contaminated by washing in infected water or from being handled by infected persons. Educating the local population on the importance of good sanitation, good personal hygiene practices, boiling water before consumption, proper cleansing before food preparation and the washing of clothes are the most important factors in limiting the spread of *Giardia*. This will help limit the number of serious juvenile chronic infections which can have current and future consequences such as malnutrition and retarded growth.

In conclusion, the results of molecular epidemiology of *Giardia* in samples from Cambodia agreed with many previous studies, not only from Southeast Asia but also corresponding with those found in various regions around the world (Volotao *et al.*, 2007; Mohammed Mahdy *et al.*, 2009; Sprong *et al.*, 2009). Both assemblages A and B are common in Cambodia, however, assemblage B was the most predominant. It indicated that anthroponotic routes play a major role of *Giardia* transmission in Cambodia. Nevertheless, *Giardia* parasites of assemblage B demonstrated higher genetic variation than the other assemblages. Because the major transmission route in *Giardia* (83.3%) is anthroponotic, in order to lower the incidence of disease, public health policies should focus on improvements in sanitation and hygiene rather than changes in animal/meat processing.

CHAPTER FOUR

COMPARISON OF THE EVOLUTIONARY ORIGIN OF *GIARDIA* SPP. FROM MALAWI AND CAMBODIA

4.1 INTRODUCTION

Giardia duodenalis is a gut protozoan infecting humans and a wide range of animals, and it is an important cause of human gastroenteritis worldwide. However, the actual role of animals such as pets and livestock as reservoirs for human infection remains unclear. *G. duodenalis* is a highly diverse parasite and it has been categorised into eight (A to H) morphologically identical genotypic assemblages with variable host occurrence. Humans are infected by assemblages A and B, which show extensive genetic variability at the sub-assemblage level. Multi-locus sequence typing (MLST) of structural and housekeeping genes is a useful approach to detect potentially zoonotic genotypes of this pathogen. Genotyping of human isolates from different locations in conjunction with epidemiological data, can provide information about potential transmission routes and epidemiological differences between *G. duodenalis* assemblages.

As described earlier (Chapter 2 and Chapter 3), sequencing has been undertaken to confirm the identity of *Giardia* isolates (*tpi*, *bg* and *gdh* markers). For more than twenty years, researchers have developed molecular biological methods to detect and characterise the parasite species, assemblages and sub-assemblages (Xiao and Ryan, 2004; Caccio, 2005; Xiao and Feng, 2008). Several studies addressing the epidemiology of *G. duodenalis* in various geographical areas using these molecular tools found that there are many assemblages and sub-assemblages of *G. duodenalis* such as human-specific, animal-specific and zoonotic species. The molecular techniques have been commonly applied to identify *G. duodenalis* assemblages and sub-assemblages in humans, animals and water samples.

In developing countries, there are some published papers describing the molecular epidemiology of *G. duodenalis* isolates in symptomatic and non-symptomatic cases based on multilocus genes and other molecular techniques (Tungtrongchitr *et al.*, 2010; Lim *et al.*, 2011; Huey *et al.*, 2013; Fahmy

et al., 2015; Anim Baidoo *et al.*, 2016; Mbae *et al.*, 2016; Moore *et al.*, 2016; Hussein *et al.*, 2016). There are few reports studied molecular analysis of the 18S rRNA gene of *Cryptosporidium* parasites from humans living in various countries (Kenya, Malawi, Brazil, the United Kingdom, and Vietnam (Gatei *et al.*, 2003) and another published paper studied on *Cryptosporidium* parasites isolated from humans living in Iran, Malawi, Nigeria, the United Kingdom, and Vietnam using a multi-locus study (Graffari and Galantari, 2014). However, there is no any report comparing *G. duodenalis* assemblages and sub-assemblages from different geographical areas by using multiple biological tools.

Therefore, the current study has investigated the use of *tpi*, *bg*, and *gdh* genes in *G. duodenalis* for characterisation of *G. duodenalis* isolated from Malawi and Cambodia. It also assessed the degree of variation in sequences between the isolates within each assemblage and their effect on the application of this gene target to a phylogenetic analysis of *G. duodenalis*.

4.2 AIMS OF THE STUDY

The aims of this study were:

1. To compare several different molecular methods by using multilocus genes for characterisation of *G. duodenalis* assemblages and sub-assemblages isolated from Malawi and Cambodia
2. To describe the genetic diversity and evolutionary relationships among *G. duodenalis* assemblages and sub-assemblages isolated from Malawi and Cambodia

The biological questions addressed in this study are: Are there different percentages of the outcomes by using different multilocus genes for characterisation of *G. duodenalis* assemblages and sub-assemblages isolated from Malawi and Cambodia?

1. Are there genetic difference (genetic variation or genetic diversity) between *G. duodenalis* in Malawi and Cambodia?
2. Are there different distributions of *Giardia* assemblages/sub-assemblages and infections in different geographical locations between Malawi and Cambodia?

4.3 MATERIALS AND METHODS

4.3.1 Isolates

Isolates used in this study comprised those which were successfully sequenced in Chapter 2 and Chapter 3. There were 89 isolates (Malawi = 49 samples and Cambodia = 40 samples) which were analysed by *tpi* gene. 33 isolates (Malawi = 14 samples and Cambodia = 19 samples) were chosen for making a phylogenetic tree of the *bg* gene sequence. For the *gdh* gene, there were 83 isolates (Malawi = 40 samples and Cambodia = 43 samples) which were used for this study.

4.3.2 Purification of the secondary PCR product

PCR products were generated as described in Chapter 2. PCR products were purified using the QIAquick® PCR Purification kit (QIAGEN®) as per the manufacturer's instructions, and were sent for sequencing to the Core Genomic Facility, Medical School, University of Sheffield, UK. Products were sequenced in both directions (forward and reverse) using an Applied Biosystems® 3730 DNA Analyser.

4.3.3 Preparation of DNA sequences

The chromatograms were checked for the quality of sequencing by manual methods before drawing phylogenetic trees. The sequences were annotated by using BioEdit ver. 7.2.0 (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). Editing and assembling in sequence contigs was then undertaken. BioEdit ver. 7.2.0 also was utilised for cutting and trimming of sequences.

4.3.4 Blast and alignment

The Basic Local Alignment Search Tool (BLAST) (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was performed to check sequence contigs of *G. duodenalis* from each isolate from Malawi and Cambodia and to confirm the isolates from various regions belonged to *G. duodenalis*. Clustal W programme was used for multiple sequences alignment.

4.3.5 Evolutionary informatics

Phylogenetic trees are used to demonstrate evolutionary relationships. Nodes represent the taxonomic unit, branches indicate the descent and ancestry of the unit. When using data from nucleotide sequences, there are four common ways of drawing phylogenetic trees.

1. Distance matrix method – here the number of nucleotide substitutions between sequences are computed for all pairs of taxa and a phylogenetic tree is constructed by using an algorithm based on the functional relationships among distance values.
2. Maximum parsimony (MP) – here the character states (nucleotide at a site) are used and the shortest pathway leading to these character states is chosen as the best tree.
3. Neighbour joining (NJ) – here the nearest sequential neighbour is found (using nucleotide homology) that minimises the length of the tree when joined.
4. Maximum likelihood (ML) – here the pattern of nucleotide differences at a site among all sequences involved is assessed and the one with the ‘largest maximum likelihood’ preferred (Li, 1997). ML trees was used due to it is the best combination of evolutionary model and rate variation among sites.

Maximum parsimony (MP) is popular because it is a relatively simple method. However, it is difficult to create the most parsimonious tree because there is no algorithm available to do this. MP trees are always the shortest ones, but this means that the most accurate tree which more correctly denotes evolutionary history, can be underestimated by the MP method. Therefore, the maximum parsimony tree can be an incorrect tree to use, especially if sequences have evolved quickly and/or at different speeds in different parts of the tree. MP can also be viewed as statistically inconsistent. Joe Felsenstein in 1978 demonstrated that the method cannot be relied upon to give consistent results when under “long branch attraction” methods (Felsenstein, 1978).

Maximum likelihood (ML) methods use a model of sequence evolution to work out the likelihood that a given alignment evolved on a tree topology. Searching over the set of possible tree topologies (not exhaustively), the method chooses the one that gives the highest likelihood of the observed data. Estimations must also be made from the existing data from given criterion such as substitution rates, variability in the rate of evolution, and edge lengths.

In this Chapter, I used the Maximum likelihood (ML) method. MP assumes that trees with the minimum number of evolutionary changes are the most preferable. MP is based on the number of character-state changes to construct all possible trees and give each a score. Advantages of MP are that they can reflect the ancestral relationship, that they can be used for morphological data, they use all known evolutionary information and are faster than the ML method. The disadvantages of MP are that it gives limited data about branch length, requires considerable computation time and the tree can be yield biased under certain conditions. It is also unreliable, especially when the rate of evolution is fast or variable. ML uses a statistical tool so that it constructs all possible trees of evolutionary history from an observed data set. Advantages of ML are that it is more accurate than other methods, it is recommended for molecular data, it is often used to test an existing tree, that all the sequence information is used, it evaluates all possible trees and sampling errors have the least effect on the method. Disadvantages of ML are that it can be extremely slow and impractical for analysing large sets of data. Therefore, ML allow greater use of available information and greater accuracy; hence this method was used for construction of the phylogeny in this work.

Bootstrap, a computational technique, is used to estimate the statistical validity of the tree; this gives a confidence level of the statistical hypothesis (Felsenstein, 1985).

The phylogenetic tree analysis were built using MEGA ver. 7 after determining the optimal nucleotide substitution model.

4.3.6 Phylogenetic tree construction

Sequences were imported into the Molecular Evolutionary Genetics Analysis (MEGA) program ver. 7 (Tamura *et al.*, 2013) and aligned using ClustalW. The sequences were trimmed to be the same length by editing and cutting the excess nucleotides from each isolate.

To identify the assemblages and sub-assemblages of *G. duodenalis*, sequences successfully sequenced in both directions (without overlapping nucleotides) were aligned and compared with reference sequences from the major *G. duodenalis* sub-assemblages AI, AII, BIII, BIV and another group of B (B unidentified). All representative sequences retrieved from the GenBank are demonstrated in Table 4.1.

Table 4.1 Details of reference DNA sequences from GenBank

GenBank Accession number	Loci (gene)	<i>Giardia</i> assemblages	Isolates	Reference
L02120	<i>tpi</i>	AI	WB	Mowatt <i>et al.</i> , 1994
U57897	<i>tpi</i>	AII	-	Baruch <i>et al.</i> , 1996
DQ650648	<i>tpi</i>	AIII	-	Lalle <i>et al.</i> , 2008
AF069561	<i>tpi</i>	BIII	BAH-12	Monis <i>et al.</i> , 2002
AF069560	<i>tpi</i>	BIV	AD-19	Monis <i>et al.</i> , 2002
HM140722	<i>tpi</i>	Sweh171	-	Lebbad <i>et al.</i> , 2010
KC6325558	<i>tpi</i>	HS114	-	Koehler <i>et al.</i> , 2013
KC632557	<i>tpi</i>	HS29	-	Koehler <i>et al.</i> , 2013
HM140714	<i>tpi</i>	Sweh060	-	Lebbad <i>et al.</i> , 2010
GU564279	<i>tpi</i>	RM1	-	Wang <i>et al.</i> , 2010
EF685684	<i>gdh</i>	BIII	BAH-12c14	Lasek <i>et al.</i> , 2007
AF069059	<i>gdh</i>	BII	BAH-12	Monis <i>et al.</i> , 2002
AY178738	<i>gdh</i>	BIV	`-	Monis <i>et al.</i> , 2002
L40548	<i>gdh</i>	BIV	AD-7	Monis <i>et al.</i> , 2002
HM165216	<i>bg</i>	B1-5	Sweh042	Lebbad <i>et al.</i> , 2010
HM165208	<i>bg</i>	B1-3	Sweh001	Lebbad <i>et al.</i> , 2010
AY072725	<i>bg</i>	BI	Nij-5	Caccio <i>et al.</i> , 2002
HM165214	<i>bg</i>	B1-1	Sweh042	Lebbad <i>et al.</i> , 2010
HM165213	<i>bg</i>	B1-2	Sweh033	Lebbad <i>et al.</i> , 2010
BAH8	<i>bg</i>	BIII	BAH8	Caccio <i>et al.</i> , 2002
AY072728	<i>bg</i>	BIV	ISSGF-4	Caccio <i>et al.</i> , 2002
HM165226	<i>bg</i>	Sweh198	-	Lebbad <i>et al.</i> , 2010

MEGA ver. 7 was used to create maximum likelihood trees after deciding on the best nucleotide substitution model, such as by combining the evolutionary model and the rate variation among sites. Genbank was used to compare reference sequences and then BLAST was used to determine whether they matched other sequences known to be representative *Giardia* assemblages and/or *Giardia* sub-assemblages. Also other species of *Giardia*, *G. ardeae* (Genbank accession number AF069060 for *gdh* gene and AF069564 for *tpi* gene) and *G. muris* isolate Swerat157 bg gene) were used as a root of phylogenetic trees (outgroup).

If a sequence showed one or more differences upon comparison with referenced sequences, they were checked with BLAST to see if they matched sequences previously deposited. Sequences deemed to be best matches were then used and incorporated into the final ML tree. This tree was constructed using one sequence to represent each differing genotype identified using 1000 replications. Sequences that could only be sequenced one way were used but kept separate from the sequences showing two-way direction.

For the sake of consistency with previous *Giardia* studies, the following rules and terminology were observed. Single nucleotide polymorphisms (SNPs) observed in the sequences were distinguished between novel polymorphisms and heterogeneous positions. Heterogeneous positional sequences were aligned directly with the representative sequences of the genotypes confirmed by sequencing in both directions and the heterogeneous positions in the sequence were indicated using the IUPAC nucleotide ambiguity codes. Locus-specific subtypes were named after their closest reference sequence or isolate found in GenBank such as AI, AII, Sweh198, followed by a lower case letter such as AIa, if the sequence showed at least one novel polymorphism. Assemblage A multi-locus genotypes (MLGs) were named following the nomenclature proposed by Caccio *et al.* (2008) capital letter to indicate the assemblage followed by a Roman numeral for the sub-assemblage and an Arabic numeral for the MLG such as AI-1, AI-2.

4.4 RESULTS

4.4.1 *G. duodenalis* sub-assemblages and genotypes by locus

4.4.1.1 *Triose phosphate isomerase* gene

Phylogenetic analysis of the *tpi* gene using the Maximum likelihood method showed that *G. duodenalis* isolates formed four different clades (Figure 4.1). The phylogenetic position of all *G. duodenalis* isolates was consistent with their preliminary classification.

Phylogenetic relationships of the isolates from different locations were assessed by using the ML phylogram method (Figure 4.1). The findings showed two *G. duodenalis* assemblage A groups consisting of sub-assemblage AI and sub-assemblage AII and demonstrated five different *G. duodenalis* assemblage B groups including BIII, BIV Ad-19 (AF069560), BIV (VB960855), B (Heterogenous) and Other B (unidentified) (Table 4.2).

Among the *G. duodenalis* assemblage A, nineteen isolates (38.8%) from Malawi with sub-assemblage AII showed 100% sequence identity to a reference strain with accession number of KR260616 from Egypt. Eleven isolates (27.5%) from Cambodia with sub-assemblage AII showed 100% sequence identity to the same reference strain with accession number of KR260616. One isolate (2.5%) from Cambodia with sub-assemblage AII showed 100% sequence identity to a reference strain with accession number of U57897 from United State of America. For sub-assemblage AI, one isolate from Cambodia showed 99% sequence identity to a reference strain with accession number of KF963567.1 from Belgium.

Among the *G. duodenalis* assemblage B, 26 isolates (53.1%) from Malawi with sub-assemblage BIII showed 100% sequence identity to a reference strain with accession number of AF069561 from Australia. Twelve isolates (30%) from Cambodia with sub-assemblage BIII showed 100% sequence identity to a reference strain with accession number of KC632557 from Australia and two isolate (4.0%) from Malawi with sub-assemblage BIV Ad-19 showed 100% sequence identity to a reference strain with accession number of AF069560 from Australia. For sub-assemblage BIV (VB960855), one isolate (2.0%) from Malawi showed 100% sequence identity to a reference strain with accession number of KM190840 from Canada, twelve isolates (30%) from Cambodia sub-assemblage BIV (VB960855) had 100%, 100% and 99% sequence identity to the published strain with accession number of KM190840 (Canada), KC632562 from Australia and KP687792 from Canada, respectively. For

sub-assembly B (Heterogenous), three isolates (7.5%) from Cambodia showed 99% sequence identity to a reference strain with accession number of GU564279 from China. For sub-assembly other B, one isolate (2.0%) from Malawi and one isolate (2.5%) from Cambodia showed 100% sequence identity to a reference strain with accession number of KF843920 from Germany.

From both countries by using the *tpi* gene, Assembly A was found in 34.8% (31/89) which belonged to sub-assembly AII 33.7% (30/89) and sub-assembly AI 1.1% (1/89). Whilst Assembly B was found in 65.2% (58/89), 42.7% (38/89) belonged to sub-assemblies BIII, 14.7% (13/89) belonged to sub-assembly BIV (VB906855), 3.4% (3/89) harboured to sub-assembly B (heterogenous) and 2.3% (2/89) displayed sub-assembly other B.

Table 4.2 Summary of the 89 isolates successfully sequenced at the *tpi* gene

Sub-assembly	No. of samples from Malawi	No. of samples from Cambodia	Samples ID	References from Genbank Accession no. (% Identity) to closest references)	Hosts	Total Percentage of <i>Giardia</i> (%)
AII	19 (38.8%)	11 (27.5%)	K94, K110, K121, K122, L5, L21, L32, L33, L38, L41, L54, L78, L81, M64, M67, M87, M107, M135, M137, A28, A155, A79, A888, A1074, A19, A272	KR260616 (100 %) U57897 (100 %)	<i>Homo sapiens</i>	30 (33.7%)
AI	0 (0%)	1 (2.5%)	A530	KF963567.1 (99 %)	-	1 (1.1%)
BIII	26 (53.1%)	12 (30%)	M64, M67, M76, M79, M107, M135, M144, M137, L2, L4, L21, L27, L32, L38, L41, L42, L54, L63, L78, K29, K87, K90, K94, K97, K108, K110, B56, B68, B75, B79,	AF069561 (100 %) KC632557 (100 %)	BAH-12 HS29	38 (42.7%)

			B84, B110, B530, B702, B749, B780, B802, B1001			
BIV Ad- 19 (AF0695 60)	2 (4.1%)	0 (0%)	L33, K121	AF069560 (100 %)	Ad-19	2 (2.3%)
BIV (VB9608 55)	1 (2.0%)	12 (30%)	BK111, B10, B15, B42, B67, B102, B121, B257, B549, B747, B761, B977, B1011	KM190840 (100 %) KC632562 (100 %) KP687792 (99 %)	<i>H. sapiens</i>	13 (14.7%)
B (Hetero genous)	0 (0%)	3 (7.5%)	B84, B780, B1001	GU564279 (99%)	RM1	3 (3.4%)
Other B	1 (2.0%)	1 (2.5%)	B82, K111	KF843920 (100 %)	VB960855	2 (2.3%)
Total	A = 19 (38.8%) B = 30 (61.2%) All =49 (100%)	A = 12 (30%) B = 28 (70%) All =40 (100%)	-	-	-	A = 31 (34.8%) B = 58 (65.2%) All = 89 (100%)

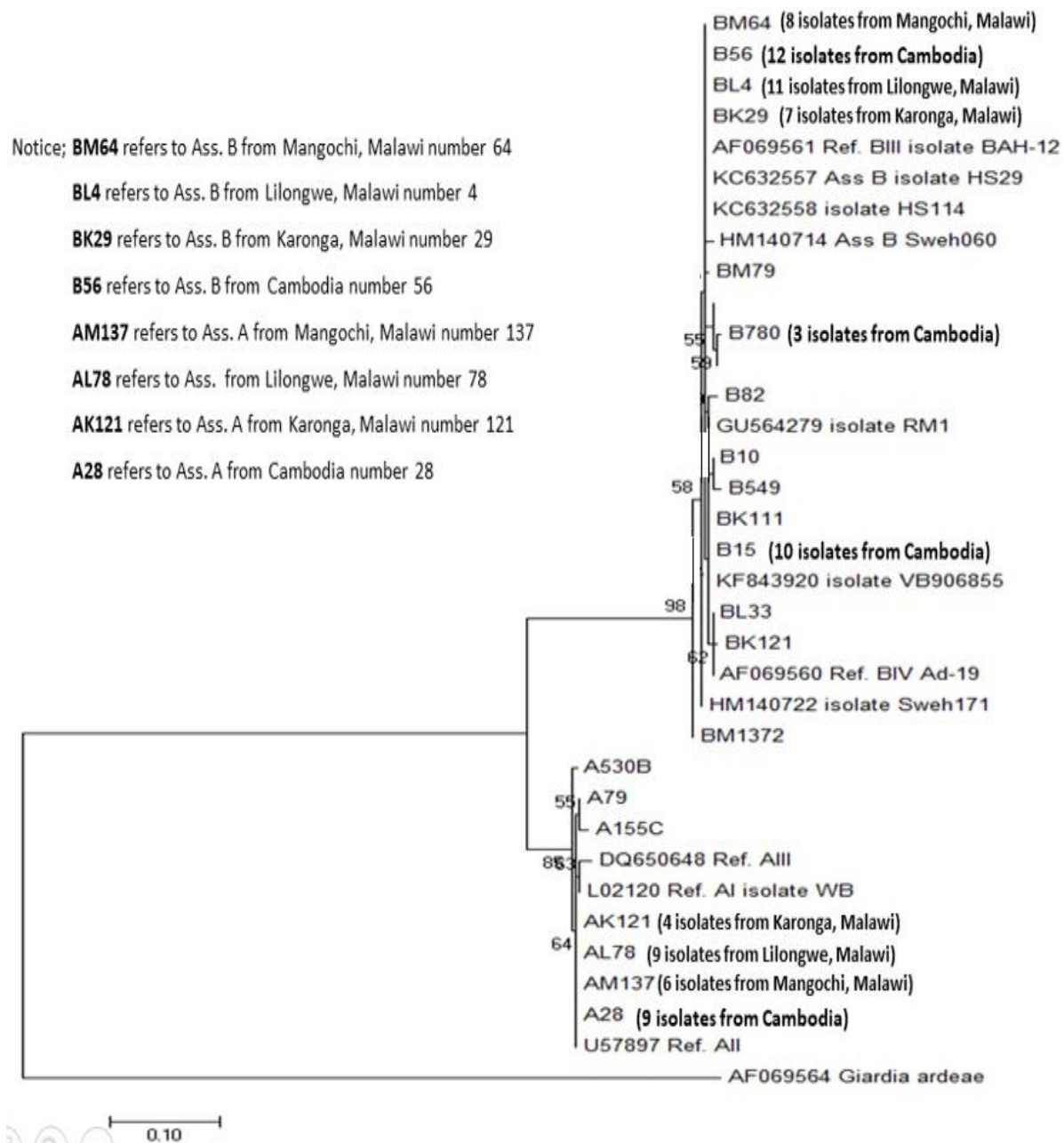


Figure 4.1 Phylogenetic relationships of *G. duodenalis* isolates showing the maximum likelihood tree of the *tpi* gene sequence of 89 isolates (Malawi =49 samples and Cambodia =40 samples). The numbers on branches are bootstrap values and the scale bar shows an evolutionary distance of 0.10 nucleotides per position in the sequence. Number above or below branches show bootstrap value (%). One representative sequence from each identified genotype is demonstrated with the total number of isolates sharing the same sequence reported in parentheses. The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates. The accession number obtained from GenBank are inserted as the reference sequences. Optimal nucleotide substitution model: Tamura-Nei parameter with gamma distribution and invariant sites. Estimated from 1,000 resampling of the sequence data.

4.4.1.2 *Beta-giardin* gene

Phylogenetic analysis of the *bg* gene using the Maximum likelihood method showed that *G. duodenalis* isolates formed four different clades (Figure 4.2). The phylogenetic positions of all *G. duodenalis* isolates were consistent with their preliminary classification.

Phylogenetic relationships of the isolates from the different location areas were assessed by analysis using the ML phylogram method (Figure 4.2). The findings showed two *G. duodenalis* assemblage A groups consisting of sub-assemblage AI and sub-assemblage AII, and demonstrated three different *G. duodenalis* assemblage B groups including BIII, B (Heterogenous) and Other B (Table 4.3). Among the *G. duodenalis* assemblage A, one isolate (7.1%) from Malawi and 2 isolates (10.5%) from Cambodia with sub-assemblage AII showed 100% sequence identity to a reference strain with accession number of JX898210 from China and KM190698 from Canada. For sub-assemblage AI, two isolate (14.3%) from Malawi and two isolates (10.5%) from Cambodia showed 100% sequence identity to a reference strain with accession number of KM190690. Among the *G. duodenalis* assemblage B, three isolates (21.4%) from Malawi with sub-assemblage BIII showed 100% sequence identity to a reference strain with accession number of JF918485 from India. Six isolates (31.6%) from Cambodia with sub-assemblage BIII showed 100% sequence identity to a reference strain with accession number of JF918500. For sub-assemblage B (Heterogenous), five isolates (35.7%) from Malawi and six isolates (31.6%) from Cambodia demonstrated 100% sequence identity to a reference strain with accession number of JF918489 from India. For sub-assemblage other B, three isolates (21.4%) from Malawi and three isolates (15.8%) from Cambodia showed 99-100% sequence identity to a reference strain with accession number of KC632652 from Thailand and JX994239.1 from China.

From both countries by using the *bg* gene, Assemblage A was found in 21.2% (7/33), 9.1% (3/33) belonged to sub-assemblage AII and 2.1% (4/33) detected to sub-assemblage AI.

Whilst Assemblage B was found in 78.8% (26/33), 27.3% (9/33) belonged to sub-assemblages BIII, 33.3% (11/33) harboured to sub-assemblage B (heterogenous) and 18.2% (6/33) displayed to sub-assemblage other B.

Table 4.3 Summary of the 33 isolates successfully sequenced at the *bg* gene

Sub-assemblage	No. of samples from Malawi	No. of samples from Cambodia	Samples ID	References from Genbank Accession no. (% Identity) to closest references)	Isolates	Total Percentage of <i>Giardia</i> (%)
All	1 (7.1%)	2 (10.5%)	L78, BG79, BG843	KM190698 (100%)	SJ	3 (9.1%)
AI	2 (14.3%)	2 (10.5%)	K10, K111, BG155, BG411	KM190690 (100%)	DG	4 (12.1%)
BIII	3 (21.4%)	6 (31.6%)	L72, K90, K135 BG10, BG27, BG82, BG817, BG977, BG989	JF918485 (100%) JF918500 (100%) JF918500 (100%)	D30 D60 D60	9 (27.3%)
B (Heterogeneous)	5 (35.7%)	6 (31.6%)	L55, , K97, K121, K110, K184, BG42, BG84, BG761, BG1001, BG1032	JF918489 (100%)	D35	11(33.3%)
Other B	3 (21.4%)	3 (15.8%)	M67, K94, L63, BG84, BG257, BG1011	KC632652 (99%) JX994239.1 (100%)	B-sho2 ECUST4064	6 (18.2%)
Total	A = 3 (21.43%) B = 11 (78.%) All = 14 (100%)	A = 4 (21.05%) B = 15 (78.9%) All =19 (100%)	-	-	-	A = 7 (21.21%) B = 26 (78.8%) All = 33 (100%)

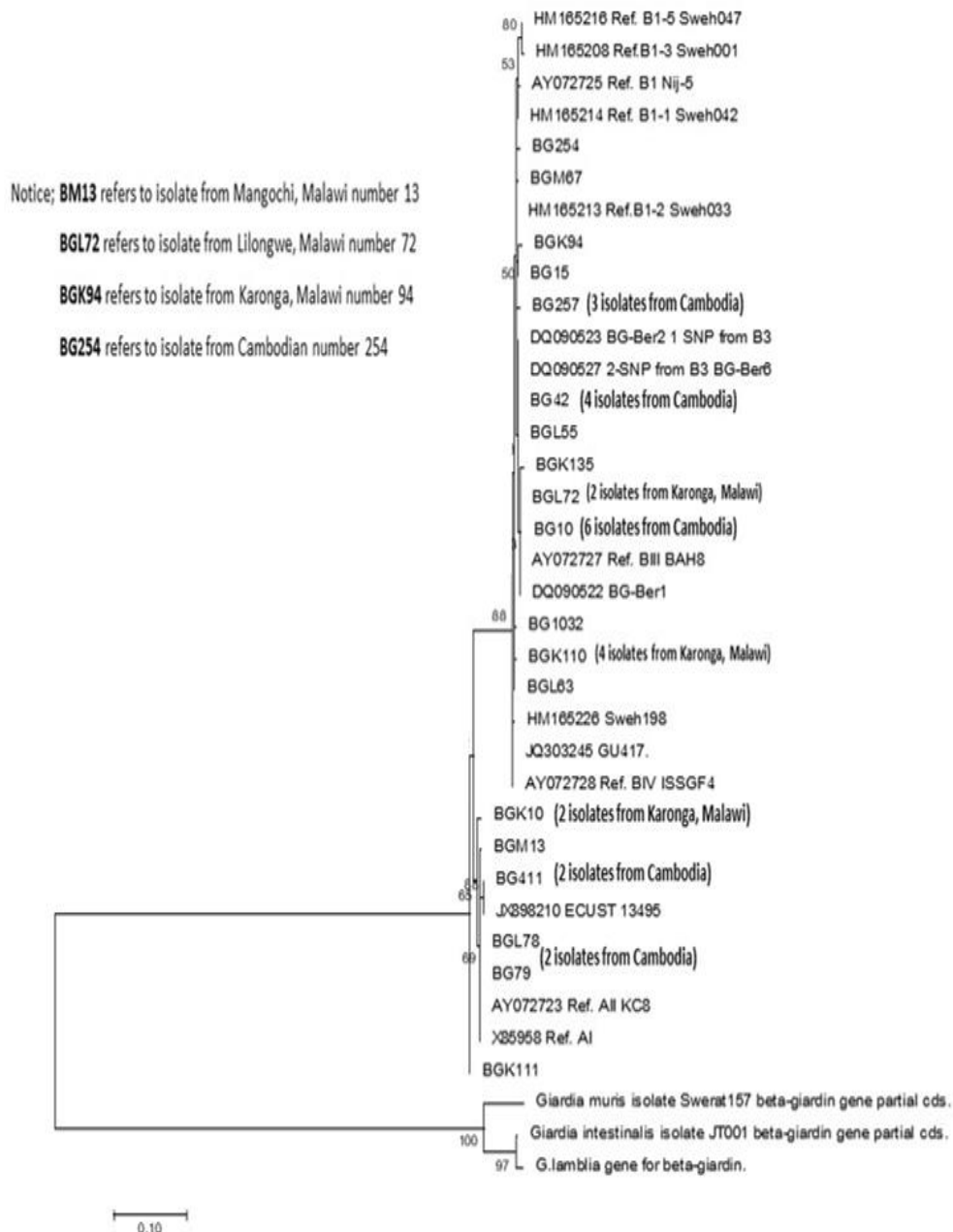


Figure 4.2 Phylogenetic relationships of *G. duodenalis* isolates showing the maximum likelihood tree of the *bg* gene sequence of 33 isolates (Malawi =14 samples and Cambodia = 19 samples). The numbers on branches are bootstrap values and the scale bar shows an evolutionary distance of 0.10 nucleotides per position in the sequence. Number above or below branches show bootstrap value (%). One representative sequence from each identified genotype is demonstrated with the total number of isolates sharing the same sequence reported in parentheses. The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates. The accession number obtained from GenBank are inserted as the reference sequences. Optimal nucleotide substitution model: Tamura-Nei parameter with gamma distribution and invariant sites. Estimated from 1,000 resampling of the sequence data.

4.4.1.3 *Glutamate dehydrogenase gene*

Phylogenetic analysis of the *gdh* gene using of the Maximum likelihood method showed that *G. duodenalis* isolates formed four different clades (Figure 4.3). The phylogenetic position of all *G. duodenalis* isolates were consistent with their preliminary classification.

Phylogenetic relationships of the isolates from the different locations were assessed by analysis using the ML phylogram method (Figure 4.3). The findings showed two *G. duodenalis* assemblage A groups consisting of sub-assemblage AI and sub-assemblage AII and demonstrated five different *G. duodenalis* assemblage B groups including BIV (KC96064.1), B (VAN/90/UBC/54), BIV (H 43), BIV (Cla145) and B (VAN/94/UBC/122) (Table 4.4). Among the *G. duodenalis* assemblage A, six isolates (13.9%, 6/43) from Cambodia with sub-assemblage AII showed 100% sequence identity to a reference strain with accession number of AY178737 (Bris-136) from Australia. For sub-assemblage AI, one isolate (2.3%, 1/43) from Cambodia showed 99% sequence identity to a reference strain with accession number of M84604 (Portland 1 strain).

Among the *G. duodenalis* assemblage B, twelve isolates (30%) from Malawi with sub-assemblage BIV (KC96064.1) showed 100% sequence identity to a reference strain with accession number of KC96064 AF069561 (BAH-12) from Australia. Thirteen isolates (30.2%) from Cambodia with sub-assemblage BIV (KC96064.1) showed 100% sequence identity to a reference strain with accession number of KC96064. For sub-assemblage B (VAN/90/UBC/54), six isolate (13.9%) from Cambodia showed 99% sequence identity to a reference strain with accession number of KP687770 from Canada (Creston). For sub-assemblage BIV (H43), 23 isolates (57.5%) from Malawi and 13 isolates (30.2%) from Cambodia showed 99% sequence identity to a reference strain with accession number of EF507682 from Brazil.

For sub-assemblage BIV (Cla145), three isolates (7.5%) from Malawi showed 100% sequence identity to a reference strain with accession number of HM134212.1 from Brazil. For sub-assemblage B (VAN/94/UBC/122), two isolates (5%) from Malawi and four isolates (9.3%) from Cambodia showed 100% sequence identity to a reference strain with accession number of VAN/94/UBC/122.

From both countries by using the *gdh* gene, assemblage A was found in 8.4 % (7/83). Sub-assemblage AII was found in 7.2% (6/83) and 1.2% (1/83) belonged to sub-assemblage AI. Whilst Assemblage B was found in 91.6% (76/83), BIV (H43)(Brazil) was the most prevalent (31.3%, 26/83), followed by sub-

assemblage BIV (KC96064.1) which was found in 30.1.-% (25/83) whereas sub-assemblage B (VAN/90/UBC/54) and sub-assemblage B (VAN/94/UBC/122) was found in the same number of cases (7.2%, 6/83), 3.6% (3/83) harboured to sub-assemblage BIV (Cla145).

Table 4.4 Summary of the 83 isolates successfully sequenced at the *gdh* gene

Sub-assemblage	No. of samples from Malawi	No. of samples from Cambodia	Samples ID	References from Genbank Accession no. (% Identity) to closest references)	Total Percentage of <i>Giardia</i> (%)
All	0 (0%)	6 (13.9%)	GD79, GD272, GD530, GD824, GD884, GD888	Bris-136 AY178737 (100%)	6 (7.2%)
AI	0 (0%)	1 (2.3%)	GD155	M84604 (100%)	1 (1.2%)
BIV (KC96064.1)	12 (30%)	13 (30.2%)	M30, M76, L5, L32, L33, L54, L55, L61, L63, L72, K29, K111 GD257, GD275, GD339, GD467, GD549, GD656, GD658, GD669, GD747, GD749, GD882, GD989, GD1001	KC96064.1 (100%)	25 (30.1%)
B (VAN/90/UBC/54)	0 (0%)	6 (13.95%)	GD67, GD84, GD121, GD591, GD738, GD1168	KP687770 (99%)	6 (7.2%)
BIV H43	23(57.5%)	13 (30.2%)	K87, K94, K135, K121, M10, M13, M35, M64, M67, M79, M107, M144, , L2, L4, L5, L32, L33, L42, L54, L55, L63, L72, L81 GD56, GD371, GD439, GD558, GD612, GD702, GD907, GD1011 GD15, GD82, GD293, GD714, GD987	EF507682 (99%)	26 (31.3%)
BIV(Cla 145)	3 (7.5%)	0 (0%)	L81, L72, M53	HM134212.1 (100%)	3 (3.6%)
B (VAN/94/UBC/122)	2 (5%)	4 (9.3%)	GD42, GD174, GD727, GD761, GD997, GD1134	VAN/94/UBC/122 (100%)	6 (7.2%)
Total	A = 0 (0%) B = 40 (100%) All = 40 (100%)	A = 7 (16.3%) B = 36 (83.72%) All =43(100%)	-	-	A = 7 (8.4 %) B = 76 (91.6%) All = 83 (100%)

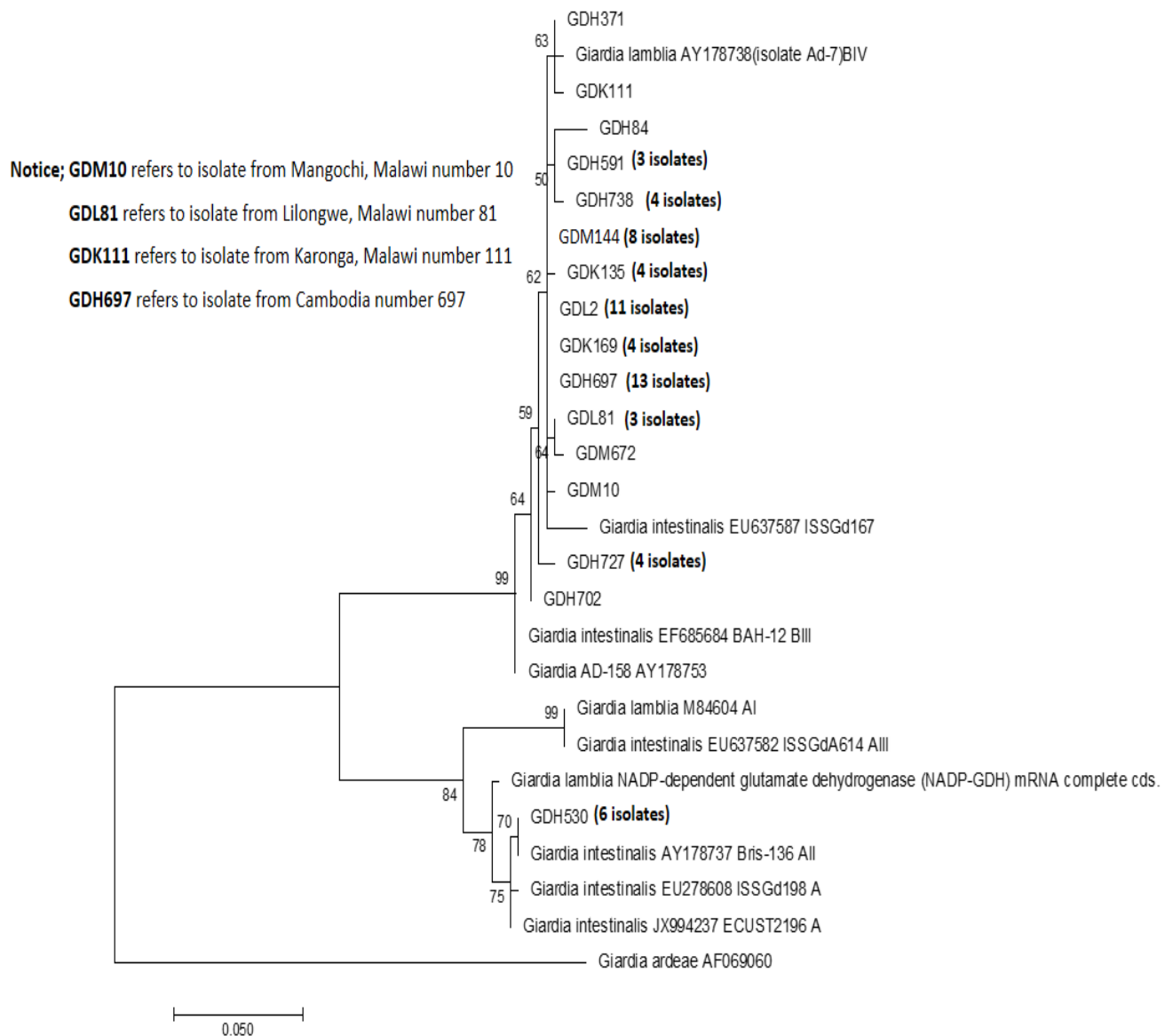


Figure 4.3 Phylogenetic relationships of *G. duodenalis* isolates showing the maximum likelihood tree of the *gdh* gene sequence of 83 isolates (Malawi = 40 samples and Cambodia =43 samples). The numbers on branches are bootstrap values and the scale bar shows an evolutionary distance of 0.05 nucleotides per position in the sequence. Number above or below branches show bootstrap value (%). One representative sequence from each identified genotype is demonstrated with the total number of isolates sharing the same sequence reported in parentheses. The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates. The accession number obtained from GenBank are inserted as the reference sequences. Optimal nucleotide substitution model: Kimura-2 parameter with gamma distribution and invariant sites. Estimated from 1,000 resampling of the sequence data.

4.5 DISCUSSION

Genotyping of *G. duodenalis* is crucial to understand the taxonomy, biology, pathology, host susceptibility and epidemiology of this parasite. For example, previous reports suggested that sub-assembly AI has a wide host range such as dogs, cats, cattle, sheep and wildlife, whereas sub-assembly AII is more common in humans than animals. Thus it is considered that sub-assembly AI is largely a zoonotic disease (transmissible from animals to humans) whilst sub-assembly AII is largely an anthroponotic disease (human disease that can be transmitted to animals) (Laishram *et al.*, 2009). In this Chapter, using isolates from Malawi and Cambodia described in Chapter 2 and 3, further characterisation of the *G. duodenalis* assembly A and B into sub-assemblies was undertaken.

The current study employed multilocus genotyping of *G. duodenalis*; four genetic markers (ssu-rRNA, *tpi*, *bg* and *gdh*) were used in Chapter 2 and as a marker, ssu-rRNA were used only for the Chapter 2 as this gene is widely used for screening of *Giardia* and enables molecular genotyping at the assembly level. Using genotyping by *bg* and *gdh*, *G. duodenalis* can be distinguished to sub-assembly level. For *tpi* gene, the *tpi* PCR-RFLP method was more sensitive compared with other genes (*bg* and *gdh*) as it could resolve the problems about mixed infections (A plus B) whereas the *bg* and *gdh* PCR-RFLP could identify these specimens as assembly A or B only (A or B alone) (Laishram *et al.*, 2012). Therefore, I used three markers (*tpi*, *bg*, and *gdh*) to investigate *Giardia* assemblies and/or sub-assemblies and explored the genetic diversity of *G. duodenalis* from Malawi and Cambodia by making phylogenetic trees.

In this study, all the human isolates classified as assembly A displayed 99-100% matches to the reference sequences, as shown by sequencing from 3 markers (*tpi*, *bg* and *gdh*).

Based on the *tpi* gene, sequence analysis demonstrated that one isolate (3.2%) belonged to sub-assembly AI, and 30 isolates (96.77%, 30/31) belonged to sub-assembly AII. By sequencing *tpi*, all instances of assembly A were typed as AII whereas upon sequencing *bg*, only one isolate (BG155) was classified as AI. For assembly B by *tpi* gene, out of 53 isolates, 38 isolates (71.7%) detected to sub-assembly BIII, 15 isolates (28.3.9%) belonged to sub-assembly BIV plus sub-assembly BIV (VB960855) and five isolates (9.4%) belonged to sub-assembly B (heterogenous)

plus a sub-assemblage other B (unidentified). All examples of assemblage B were typed as BIV (VB960855) whereas sequencing *bg*, two isolates (BG10 and BG977) were classified as BIII, three isolates (BG15, BG257 and BG1011) were classified as other B and only one isolate (BG42) was classified as B (heterogenous).

Based on the *bg* gene, sequence analysis demonstrated that four isolates belonged to sub-assemblage AI, three isolates belonged to sub-assemblage AII. For assemblage B by *bg* gene (27 isolates), 11 isolates (40.7%, 11/27) belonged to sub-assemblage B (heterogenous), nine isolates (33.3%, 9/27) detected to sub-assemblage BIII and six isolates (22.2, 6/27%) belonged to sub-assemblage other B (unidentified). Interestingly, some examples of assemblage B (*bg* gene) were typed as BIII (VB960855) whereas sequencing *gdh*, two isolates (BG10 and BG977) were classified as BIV H43 (EF507682), one isolate (ID no. 989) was classified as BIV (KC96064). Otherwise, at the *bg* gene the isolates were typed as BIII (VB960855) but at the *tpi* gene, two isolates (ID no.10 and 977) were classified as BIV (VB960855).

Based on the *gdh* gene, sequence analysis demonstrated that one isolate from Cambodia (14.3%) harboured sub-assemblage AI, 6 isolates from Cambodia (85.7%) belonged to sub-assemblage AII. Assemblage A was not found in Malawi. For assemblage B by *gdh* gene (76 isolates), 26 isolates (50%, 38/76) detected to sub-assemblage BIV H43 (EF507682, Brazil). For sub-assemblage B (VAN/90/UBC/54) and B (VAN/94/UBC/122) were identified from the same number as 6 isolates and three isolates (8.6%) belonged to sub-assemblage BIV Cla145 HM134212.1, Brazil).

For assemblage B by *gdh* gene, some examples of assemblage B were typed as BIV H43 whereas sequencing *bg*, four isolates (ID no. 32, 34, 65 and 74) were classified as other B and two isolates (ID no., 65 and 74) were identified as assemblage BIV (VB960855) by *tpi* gene, another one isolate (ID no. 82) was classified as assemblage BIII by *gdh* but were typed as other B when using *tpi* gene.

For assemblage B were typed as BIV (KC96064.1) while typing with *bg* gene, one isolate (ID no. 79) was identified as other B and when using *tpi* gene, this isolate was typed as assemblage BIV (VB960855). For the last example typing with *gdh* gene, one isolate (ID no. 70) was classified as BIV (KC96064.1) but it was typed as assemblage B (other B or unidentified to sub-assemblage) when using both *bg* and *tpi* genes.

The assignment of assemblages at all three loci showed that there was assemblage swapping in some isolates from both Malawi and Cambodia. Assemblage swapping occurs when different assemblages are noticed at different loci, but in the same isolate. Assemblage swapping has also been reported by other investigators (Caccio *et al.*, 2008; Thompson and Smith, 2011; Breathnach *et al.*, 2010). It is believed that this is caused by assemblages recombining or when an infection has more than one assemblage.

Comparison of the findings of *Giardia* in Malawi and Cambodia by using the *tpi* gene. In Malawi, the highest prevalence of *Giardia* infection was sub-assemblage BIII (53.1%, 26/49), followed by sub-assemblage AII (38.8%, 19/49), sub-assemblage BIV (4.1%, 2/49), sub-assemblage B (VB960855) (2%, 1/49), respectively. For Cambodia, the highest prevalence of *Giardia* infection was sub-assemblage BIII and B (VB960855) at the same rate (30%, 12/40), followed by sub-assemblage AII (27.5%, 11/40), sub-assemblage B (heterogenous) (7.5%, 3/40), and sub-assemblage AI (2.5%, 1/40), respectively.

The most dominant *Giardia* sub-assemblage BIII was isolated from children both in Malawi and Cambodia. I observed clinical manifestations among this sub-assemblage BIII and found that patients had coinfection with *Cryptosporidium* and Rotavirus. For example, out of 10 children who came from Lilongwe (large urban hospital), 6 persons (60%, 6/10) were infected with *Cryptosporidium*, and 4 persons infected with Rotavirus (40%, 4/10). In addition, I found that *Giardia* sub-assemblage BIII was found both in rural and urban areas; I found this sub-assemblage in Mangochi (medium size rural hospital) and Karonga (small rural hospital). With reference to sub-assemblage AII and sub-assemblage I, the results showed that all cases from Malawi had diarrhoea and children from Cambodia had diarrhoea in 85.7% (6/7) and also had abdominal pain in 71.4% (5/7). Sub-assemblage BIV (VB960855) was more prevalent in Cambodia (30%, 12/40) than in Malawi (2%, 1/49). All people who were infected with this sub-assemblage had no diarrhoea but they were observed to present with abdominal pain in 83.3% (10/12). In addition, sub-assemblage B (heterogenous) could be found only in Cambodia, all cases who were infected with sub-assemblage BIV (VB960855) had no diarrhoea but 100% (3/3) were observed to present with abdominal pain. However, it was a very small number.

Comparison of the findings of *Giardia* in Malawi and Cambodia by using *bg* gene. In Malawi, the highest prevalence of *Giardia* infection was sub-assemblage B (heterogenous) (35.7%, 5/14), followed

by sub-assemblage BIII and sub-assemblage other B at the same rate (21.4%, 3/14), sub-assemblage AI (14.3%, 2/14), and sub-assemblage AII (7.1%, 1/14), respectively. For Cambodia, the highest prevalence of *Giardia* infection was sub-assemblage BIII and sub-assemblage B (heterogenous) at the same rate (31.6%, 6/19), followed by sub-assemblage other (15.8%, 3/19), and sub-assemblage AII and sub-assemblage AI at the same rate (10.5%, 2/19), respectively.

The most dominant *Giardia* sub-assemblage B (heterogenous) was isolated from children in Malawi. I observed clinical manifestations among this sub-assemblage B (heterogenous) finding that all patients with diarrhoea were from Karonga where they had presented to a small rural hospital. From the results, it showed that children who were infected with *Giardia* from this area had coinfection with *Cryptosporidium* (60%, 3/5) and Rotavirus (40%, 2/5). For example, children who came from Lilongwe (A big urban hospital). Moreover, the assemblage B (heterogenous) was also the most predominant of *Giardia* infection in Cambodia. I found 50% of children had diarrhoea (3/6) and all persons had abdominal pain (100%, 6/6). Otherwise, sub-assemblage BIII isolated from children in Cambodia was found at the same rate as assemblage B (heterogenous) (31.6%, 6/19) isolated from children in Cambodia, whereas assemblage BIII was found in 21.4% (3/14) isolated from patients in Malawi. Interestingly, all children from both Malawi and Cambodia who infected with sub-assemblage BIII had diarrhoea. With reference to having abdominal pain, we had information about this only in Cambodia so the result found that 66.7 % of the children had abdominal pain (4/6) but 100% (6/6). For sub-assemblage other B infection in both Malawi and Cambodia children who infected with this sub-assemblage had no diarrhoea (16.7%, 1/6) which was opposite to having symptoms of abdominal pain (83.3%, 5/6).

About sub-assemblage AII, all cases from both Cambodia had no diarrhoea. Interestingly, *Giardia* sub-assemblage AI was detected in Malawi and had coinfection with *Cryptosporidium* (50%, 1/2) and Rotavirus (50%, 1/2) which were associated with rural areas in Malawi.

Comparison of the findings of *Giardia* in Malawi and Cambodia by using *gdh* gene. In Malawi, the highest prevalence of *Giardia* infection was sub-assemblage BIV H43 (57.5%, 23/40), followed by sub-assemblage BIV (KC96064) (30%, 12/40), sub-assemblage BIV Cla145 (7.5%, 3/40), and sub-assemblage B VAN 94/UBC/122 (5%, 2/40), respectively.

For Cambodia, the highest prevalence of *Giardia* infection was sub-assemblage BIV H43 and sub-assemblage BIV (KC96064) at the same rate (30.2%, 13/43), followed by sub-assemblage B VAN 90/UBC/54 and sub-assemblage AII at the same rate (13.9%, 6/43) and sub-assemblage B VAN 94/UBC/122 (9.3%, 4/43) and sub-assemblage AI (2.3%, 1/43), respectively.

The most dominant *Giardia* sub-assemblage BIV H43 was isolated from children in Malawi. Clinical manifestations among this sub-assemblage BIV H43 were observed and found that all patients had diarrhoea and they were from Karonga where they present to a small rural hospital. One person who was infected with *Giardia* in Karonga had Rotavirus infection (25%, 1/4) and another patient had coinfection with *C. parvum* (25%, 1/4). As *C. parvum* associated with zoonosis and could be found more in rural areas than urban areas. Of the findings in Lilongwe, I noticed that the patients who were infected with this sub-assemblage had coinfection with *Cryptosporidium* (*C. hominis* 45.6%, 5/11 and *C. parvum* 27.3%, 3/11) and Rotavirus (36.4%, 4/11).

For the patients from Mangochi who had no coinfection with Rotavirus and/or *Cryptosporidium*. For isolates from Cambodia, children who infected with sub-assemblage BIV H43 also had diarrhoea in 53.8% (7/13) of cases and they had abdominal pain in 84.6% (11/13).

According to infection with *Giardia* sub-assemblage BIV (KC96064), all patients from Malawi had diarrhoea. This maybe due to the children having coinfection with either *Cryptosporidium* or Rotavirus or that they had coinfection with both these pathogens. For children from Cambodia, I found that cases who infected with this sub-assemblage only 2 cases from 13 children had diarrhoea (15.4%, 2/13). The prevalence of infection of sub-assemblage BIV (KC96064) in asymptomatic cases (non diarrhoea cases) was 84.6%(11/13). However, children who infected with *Giardia* sub-assemblage BIV (KC96064) still had abdominal pain in 76.9% (10/13).

According to the *tpi* gene, I found that *Giardia* sub-assemblage BIII was more prevalent than sub-assemblage BIV (VB960855) and B (heterogenous) in Malawi and Cambodia. As infections with both sub-assemblages, the children had no diarrhoea. For Assemblage A, sub-assemblage AII was more pathogenic and AI was more likely asymptomatic. Moreover, sub-assemblage AII was found in urban areas while sub-assemblage BIII was found in both rural and urban areas. For *bg* gene, sub-

assemblage BIII maybe more pathogenic than sub-assemblage B (heterogenous). All cases infected with sub-assemblage BIII were observed to present with diarrhoea whereas patients infected with sub-assemblage B (heterogenous) had diarrhoea in 50% of the cases, however all cases were observed to present with abdominal pain. Otherwise, sub-assemblage B (heterogenous) and sub-assemblage AI could be found in both rural and urban areas. For *gdh* gene, sub-assemblage BIV H43 maybe more pathogenic than both sub-assemblage BIV (KC96064) and sub-assemblage B VAN 94/UBC/122 in Malawi. For Cambodia, sub-assemblage BIV H43 was more prevalent than other sub-assemblages (eg. sub-assemblage BIV (KC96064) sub-assemblage B VAN 90/UBC/54, sub-assemblage B VAN 94/UBC/122 and sub-assemblage AI and AII). Comparison of the finding of *Giardia* in both countries, sub-assemblage BIV H43 was more prevalent in Malawi (57.5%, 23/40) than in Cambodia (30.2%, 13/43) and the prevalence of BIV H53 infection was significantly different (fisher's exact test $p = 0.004$). Moreover, the sub-assemblage B VAN 90/UBC/54, sub-assemblage AII and AI were not found in Malawi. Otherwise, sub-assemblage BIV Cla145 was not found in Cambodia.

The present study was the first time that data on the prevalence and genetic diversity of *G. duodenalis* isolates in Malawi and Cambodia had been carried out. The genotyping results show that all *Giardia* infections in these population are due to *G. duodenalis* assemblages A and B. This confirms the results of a number of studies performed elsewhere (Caccio and Ryan, 2008). Distribution of different assemblages differs among and within countries, as studies in various countries have demonstrated a diverse prevalence of assemblages A and B (Feng and Xiao, 2011). These results have shown that children in urban informal settlements in Cambodia, predominantly carry *Giardia* assemblage B, which is in agreement with several reports from various areas around the world (Ignatius *et al.*, 2012; Ankarklev *et al.*, 2012; Lebbad *et al.*, 2011). The predominance of one *G. duodenalis* assemblage over another in a particular area has been attributed to both biological and geographical factors, in certain endemic areas (Volotao *et al.*, 2007). The reasons for the geographic diversity in the predominance of the *Giardia* assemblages are still not known. It may be due to differences in the dynamics of transmission. Our study identified both sub-assemblages BIII and BIV in Malawi and Cambodia, assemblage BIII with isolates from Malawi in 53.1% (26/49) and isolates from Cambodia in 30% (12/40) whereas assemblage BIV with isolates from Malawi (8.2%, 4/49) and isolates from Cambodia (40%, 16/40) by the *tpi* gene. The *bg* gene showed isolates from Malawi in 57.1% (8/14) and isolates

from Cambodia in 47.4% (9/19) whereas the *gdh* gene demonstrated assemblage BIV isolates from Malawi in 70% (28/40) and assemblage BIV isolates from Cambodia in 53.5% (23/43). The percentages of each sub-assemblage from isolates both in Malawi and Cambodia were different dependent on the markers used. In this study, the finding of the sub-assemblage BIV from Cambodian isolates were similar but for the Malawian isolates which were different from the results in Thailand by using *gdh* gene (Tungtrongchitr *et al.*, 2010). Genotyping of *G. duodenalis* was undertaken in 61 faecal specimens (with and without diarrhoea) from Bangkok and in rural areas of Thailand; assemblage B was found in 51%, mixed infection of both assemblages was detected in 41% and the assemblage A alone was detected in 8% by using the *tpi* gene. The prevalence of assemblage B and mixed infection (A/B mixed) were significantly higher than assemblage A alone. The sub-assemblage AI was found in 12%, the sub-assemblage AII was found in 88% by using the *bg* gene. Using the *gdh* gene, the sub-assemblage BIII was detected in 45%, and the sub-assemblage BIV was identified in 54.5%. Interestingly, the sub-assemblage AI was found in 100% from symptomatic cases and the sub-assemblage BIII sub-assemblages was detected in 50% (Tungtrongchitr *et al.*, 2010). In addition, most of the symptomatic cases in Bangkok (urban area) were adults and the elderly (80%) whereas the main incidences of symptomatic cases in the rural specimens were children (85%) (Tungtrongchitr *et al.*, 2010). In Africa, infection with *G. duodenalis* assemblage B, sub-assemblage BIII was more prevalent than infection with sub-assemblage BIV while this differed from findings in North-America where more infections were related with sub-assemblage BIV, and only a few with sub-assemblage BIII (Feng *et al.*, 2011). This however differs with our findings, where BIV is more prevalent. Our study however agrees with findings from Thailand where Assemblage B, sub-assemblage BIV was found to be the most common in preschool children (Boontanom *et al.*, 2011). Phylogenetic analysis of the isolates demonstrated that the assemblage B test isolates, formed two clusters. This could be attributed to genetic variation between reference sequences and the test samples, which after comparison of base pair position with the reference B assemblages revealed sequence profile variation within our isolates. A high degree of polymorphism in assemblage B has been observed in other studies (Caccio *et al.*, 2008; Lebbad *et al.*, 2011) and has been further investigated by cloning (Hussein *et al.*, 2009; Kosuwin *et al.*, 2010).

This study demonstrates the genetic diversity of *Giardia* assemblage B, and children in the study areas both in Malawi and Cambodia may have had different clinical responses to infections with different *Giardia* assemblages/sub-assemblages. Considering that *Giardia* assemblage B was the predominant species in these study populations. It can be concluded that transmission in children in both locations is predominantly anthroponotic. This study implies that the distribution of *G. duodenalis* genotypes can vary according to geographical location. The human faecal samples were obtained from children in Malawi and the DNA samples were received from patients in Cambodia, and it can be concluded that both assemblages A and B are common in both countries. Nonetheless, further study is required. Ideally, samples should also be collected from healthy donors as it would be better to do research from case and control study at the same time.

In conclusion, the study provides some preliminary data on assemblage and sub-assemblage distribution of *G. duodenalis* in Malawi and Cambodia and highlighted that *Giardia* assemblage A and B are prevalent in children in both countries, with a predominance of assemblage B. These findings suggest that anthroponotic transmission could be a dominant transmission route for giardiasis in both countries. Therefore, there is need to explore the possibility of zoonotic transmission.

CHAPTER FIVE

AN INVESTIGATION OF THE USE OF GAS CHROMATOGRAPHY MASS SPECTROPHOTOMETRY (GCMS) FOR DIAGNOSING *CRYPTOSPORIDIUM* INFECTION IN CHILDREN IN MALAWI

5.1 INTRODUCTION

Cryptosporidium is a protozoan parasite, infecting both humans and other animals. This parasite is an emerging pathogen as a common cause of diarrhoea worldwide and it is the commonest protozoal cause of acute gastroenteritis in some countries (Feng *et al.*, 2007a; Ryan *et al.*, 2004). *Cryptosporidium* can cause protracted diarrhoea in both immunocompetent and immunocompromised individuals worldwide (Farthing *et al.*, 1996; Kappus *et al.*, 2004). In immunocompetent patients, the illness is self-limiting but generally lasts one or two weeks and sometimes longer. In immunocompromised hosts such as those with HIV-AIDS the infection can be prolonged and life threatening (Kjos *et al.*, 2005). Human Cryptosporidiosis is an illness caused by *Cryptosporidium* spp., it can cause profuse watery diarrhea, abdominal discomfort, loss of appetite, nausea, vomiting, mild fever and can lead to death if the patient becomes severely dehydrated. Cryptosporidiosis is often misdiagnosed due to technical issues relating to laboratory diagnosis. Since the oocysts of this protozoa are very small and difficult to identify, this can lead to false-negative results (Verweij *et al.*, 2003). Previously, the most commonly used biological test for the presence of *Cryptosporidium* was identification of the parasite in the patient's stools using light microscopy. This method is specific but laborious and time-consuming (Behr *et al.*, 1997; Feng *et al.*, 2007a). Moreover, the procedure requires well-trained personnel for differentiation of the protozoa parasites which are difficult to identify (Verweij *et al.*, 2003). Therefore, proper identification of *Cryptosporidium* spp. requires accurate and reliable methods (Verweij *et al.*, 2003). In the past decade, immunofluorescent stains and polymerase chain reaction (PCR) tests have been developed. Both techniques are accurate, reliable, very specific and highly sensitive. However, these assays have the disadvantage of being expensive, require multiple procedural steps and highly technological

equipment as well as trained personnel (Verweij *et al.*, 2003). The parasitological diagnosis of infectious diarrhoea sometimes takes several days from the collection of the faecal sample. This may be due to transportation of the specimens to the diagnostic laboratory (Johal *et al.*, 2002). However, the main problem is time consuming conventional diagnostic methods (Verweij *et al.*, 2003).

Consequently, researchers attempted to find a new and effective method to rapidly detect gastrointestinal infections. Several groups such as doctors, nurses, patients and researchers have observed that faeces often smell different when people had diarrhoea (Garner *et al.*, 2007; Probert *et al.*, 2004). Volatile organic compounds (VOCs) are potential biomarkers for rapid diagnosis of gastrointestinal diseases (De Lacy Costello, 2008; Garner *et al.*, 2009; Garner *et al.*, 2007; Probert *et al.*, 2009; Probert *et al.*, 2004).

VOCs are a group of very small carbon-based molecules which have a high vapour pressure at room temperature due to their low boiling points, and because of this many of the molecules evaporate into the surrounding air (Garner *et al.*, 2007; Probert *et al.*, 2004). The major sources of VOCs include breath, sweat, blood, urine and faeces and can vary with diet, sex, age, physiological status and, possibly, genetic background. Little is known about the VOCs in faeces and their potential health consequences. Most scents contain a combination of VOCs and for this reason they can be used in the fragrance industry. They may also play a role in health and disease with several being efficient biomarkers of different conditions (Garner *et al.*, 2007; Garner *et al.*, 2009; Probert *et al.*, 2004; Probert *et al.*, 2009).

Previous studies have attempted to define and compare the VOCs emitted from both healthy donors and those suffering from gastrointestinal diseases. They had three main hypotheses 1) VOCs would be common in health 2) VOCs would be constant amongst individuals 3) VOCs would show specific changes amongst patients with gastrointestinal disease (Garner *et al.*, 2007). A list of VOCs in healthy individuals (Cohort study) was compiled and then compared against those VOCs emitted by patients suffering from the following diseases, ulcerative colitis, *Campylobacter jejuni*, and *Clostridium difficile*. The volatiles from the various patients faeces were collected by solid-phase microextraction and analyzed by gas chromatography/mass spectrometry (GCMS). In the cohort study, 297 volatiles were identified. Common to all samples were; ethanoic, butanoic, pentanoic acids, benzaldehyde,

ethanal, carbon disulfide, dimethyldisulfide, acetone, 2-butanone, 2,3-butanedione, 6-methyl-5-hepten-2-one, indole, and 4-methylphenol. Forty-four compounds were shared by 80% of subjects. In the longitudinal study, 292 volatiles were identified, with some *inter* and *intra* subject variations in VOC concentrations over time. Volatile patterns from healthy donors were compared to those from donors with the previously listed diseases showing that the patterns emitted were significantly different. It was hoped that with these results, further research could lead the way to the development of a rapid diagnostic device based on VOC detection.

For investigating the use of GC/MS as a novel method for studying the VOCs in the faeces of patients with *Giardia* infection in North West United Kingdom, faecal samples were collected from patients with confirmed isolated *Giardia* infection. Stool samples were also collected from patients that had diarrhoea but no identifiable infection, these samples would act as controls. Faecal gases were extracted and GC/MS used to identify any VOCs. Metab, R (Aggio *et al.*, 2011) and Metaboanalyst (Xia *et al.*, 2009) were used to analyse any significant differences between the VOCs. After analysis, it was noted that Esters, Acids and Alcohols dominated the composition of the compounds and that the compounds contained no Amides, Alicyclic, Ether or nitrogen. Further analysis of both the *Giardia* and Control groups showed that in excess of 100 VOCs were detected of which 10 were noticeably different between the two groups. Much higher levels of 2,2,4,4-tetramethyloctan, acetic acid and 2,2,4,6,6-pentamethylheptanone ($p < 0.0001$) were found in the *Giardia* samples. With reference to abundance there were 10 VOCs with noticeably different levels between the two groups the two most significant being acetone and 2-butanone which were down regulated in the *Giardia* group. The experiment showed noticeable differences in the VOC profiles of the stool samples between the *Giardia* and Control (non infected) groups. This difference can be explained metabolically as *Giardia*'s metabolism has its own distinctive metabolic pathways (Vernon *et al.*, 2015). The method gives the best results when used for uncomplicated mixtures and where the chemicals are already understood and calibrated.

The question to be asked in this study is; there will be a difference in the VOCs present in faeces of patients who are infected with *Cryptosporidium* spp. compared with the VOCs present in faeces from negative controls (without *Cryptosporidium* spp.)?

5.2. AIMS AND OBJECTIVES

The study objectives were:

1. To define and analyse VOCs emitted from the faeces of patients with gastrointestinal disease.
2. To compare the VOCs present in the *Cryptosporidium* positive faecal samples with the VOCs present in the control samples.
3. To identify any potential biomarkers of *Cryptosporidium* spp. infection.

5.3 MATERIALS AND METHODS

5.3.1 Sample Preparation

All faecal samples were provided by Prof. Nigel Cunliffe (Department of Clinical Infection, Microbiology and Immunology, University of Liverpool). The stool samples used here were collected as part of previous studies on rotavirus in Malawi (Cunliffe *et al.*, 2009a; Cunliffe *et al.*, 2009b). The details of collection, the method (methodology) and storage are the same as described in Chapter 2. After the samples were used for the experiments in Chapter 2, the volumes of the remaining faecal specimens were checked to see if they had enough volume to enable the analysis of their Volatile Organic Compounds. At least 0.1-0.75 g faeces was taken from each of 11 samples in which *Cryptosporidium* had been previously detected by Direct Fluorescent assay (DFA) and PCR. Each sample was put into 10 ml vials. Approximately 0.5-0.75g faeces were aliquoted from 11 patients who had no *Cryptosporidium* infection by DFA and PCR methods. The samples were stored in 11, 10 ml headspace vials. All of the 22 vials were kept in freezers (-80 °C) before and after use.

Table 5.1 Demographics of the study participants

Information	<i>Cryptosporidium</i> samples	Negative control (no <i>Crypto</i>)
Total no.	11	11
Sex	F=7, M=4	F=7, M=4
Age	0-1 year = 9 1-2 years = 0 2-3 years = 0 3-4 years = 2	0-1 year = 9 1-2 years = 2 2-3 years = 0 3-4 years = 0
Rotavirus infection	Rotavirus = 2 No rotavirus = 9	Rotavirus = 0 No rotavirus = 11
Sources of samples	Lilongwe = 8 Mangochi = 1 Karonga = 2	Lilongwe = 7 Mangochi = 0 Karonga = 4
Volume of faeces	100 µl = 2, 200 µl = 2 250 µl = 2, 300 µl = 1 400 µl = 1, 500 µl = 2 750 µl = 1	500 µl = 11
Consistency of faeces	Liquid = 4 solid + liquid = 1 solid = 5 water (a lot) = 1	solid + liquid = 11

5.3.2 VOCs Extraction

Solid-phase Microextraction (SPME) is a technique created by Arthur and Pawliszyn (1990). It is an extremely efficient method of preparing samples without the use of solvents. SPME has proved to be suitable for use in a wide range of chemical analysis, since its early use in the environmental and food safety areas. It works well as a technique when paired with mass spectrometry (MS). The traditional processes of liquid-liquid extraction (LLE)-extraction, concentration (derivatization) and transfer to chromatograph, can now all be performed using one device and one procedure. This enables much easier preparation of samples. A fused-silica fibre that is externally coated with the appropriate stationary phase, analytes will be directly extracted to the coating of the fibre. SPME can be used in conjunction with gas chromatography, high-performance liquid chromatography and capillary electrophoresis without restricting or compromising MS. The advantages of using SPME include reduced expenditure on purchase and disposal of solvents, speedier sample preparation and improved detection limits. SPME is well matched for MS applications making sample preparation less complicated yet giving both sensitivity and versatility in detection (Vas & Vekey, 2004).

SPME was used to extract the VOCs from each sample by a 2 cm long fused silica coated fibre. Each sample in turn was defrosted for 3 minutes and heated in a water bath at 60°C for 30 minutes. The SPME needle was injected into the vial while it remained in the water bath. The fibre was exposed for 20 minutes. The fibre was then retracted back into the device and the SPME needle removed from the sample. The sample was left to cool and then returned to the freezer (Garner *et al.*, 2007).

5.3.3 Analysis of volatile organic compounds

The samples were analysed by a Perkin Elmer Clarus 500 GC/MS quadrupole bench top system fitted with a 1 mm quartz liner. Gas within the headspace was then sampled and analysed using gas chromatography (GC) and mass spectroscopy (MS) (Busch, 2010). Gas chromatography-mass spectrometry (GCMS) was used to separate and identify the VOCs (Cerdán-Calero *et al.*, 2012). The GC column was a 60 M Zebron ZB-624, 0.25 mm ID, 1.40 µm film thickness (Phenomenex Melville House, Queens Avenue, Hurdsfield Industrial Estate, Macclesfield) (Garner *et al.*, 2009a). The carrier

gas was 99.9995% pure helium. After SPME VOCs extraction, the SPME needle was inserted into the GC injection port and the VOCs desorbed from the fibre onto the GC column by thermal desorption at 280°C (Dixon *et al.*, 2011). The fibre was then immediately transferred to the GC port for thermal desorption at 280 °C with the split valve closed throughout. The GC temperature program was operated as follows: (35 °C) 5 min, ramp of 7 °C min⁻¹ to 250 °C, and finally held at 250 °C for 12 min (total run time 41 minutes). The VOCs are then detected by MS in electron ionization (EI) mode scanning from a mass ion (ranging from 17 to 350 of ions) (Busch, 2010). The VOCs interact with the column and are separated according to their chemical properties such as molecular weight and polarity (Cerdan-Calero *et al.*, 2012).

5.3.4 AMDIS

The Automated Mass Spectral Deconvolution and Identification System (AMDIS) is a piece of easily available software that is designed to interpret the results of gas chromatography and mass spectrometry (GCMS). It has a large reference file enabling it to identify and match up with specific compounds. AMDIS's aim is to make the identification of compounds less strenuous when using GCMS while still retaining high levels of both accuracy and reliability such as those offered by more traditional methods (Gary Mallard and Reed, 2000). Although the model peak method (Droney *et al.*, 1976) had demonstrated reliability when conducting mass testing, extraction of weak signals was very often insufficient. This was because the method was unable to separate signal from noise. Current work now processes ion abundances as signal to noise units rather than absolute abundances. This allows for rational settings of thresholds throughout the extraction process. Optimised Spectrum Comparison Function (Stein and Scott, 1994) was the basis for the identification of chemicals, the use of information derived from GCMS data enabling incorporation of other data. The spectrum comparison process was further augmented by four refinements developed after data analysis of previous results. These steps are (sequentially) 1) noise analysis 2) component perception 3) spectral “deconvolution” 4) compound identification. AMDIS has several user selectable parameters to aid detection such as match factor, sensitivity, resolution and SIM versus scan analysis. AMDIS can often yield a high rate of false identifications of metabolites, referred to as the false positive rate.

This often leads to smaller portions of the chromatograph having to be analysed. However, algorithms such as Pscore are available, Pscore is a GCMS based retention time (RT) scoring algorithm used to assess the likelihood that the observed RT's in a biological sample correspond to known metabolites within a user defined spectral library.

AMDIS was used to analyse the data produced by the GCMS (Zhang *et al.*, 2006). Mass spectrums and fragment patterns were compared to a library of compounds in the National Institute of Standards and Technology (NIST) to identify a compound match for each of the compound present in each sample (Samokhin & Revelsky, 2011). Once each chromatogram was analysed and compounds in each sample identified, the results were compared. All compounds were identified by comparison with the mass spectral NIST 05 library, followed by manual visual inspection. This process was repeated for the 22 samples and each run was alternated between *Cryptosporidium* positive samples and *Cryptosporidium* negative control samples. For a control, an empty glass vial was stored under similar conditions to the sample vials and analysed for VOCs using the standard method.

5.3.5 MetaboAnalyst 2.0 programme

MetaboAnalyst is a web programme for high-throughput metabolic data processing, analysis (data normalization, statistical analysis and high-level functional interpretation) and annotation (Xia *et al.*, 2009). The design of which is to suit biologists with limited background in statistics. It can be used to compile data relating to metabolomics analysis. The process can be completely undertaken in 45 minutes, and offers detailed descriptions of results of all metabolic data. Recently developed approaches including metabolite set enrichment analysis and metabolic pathway analysis can also be undertaken using MetaboAnalyst. It is the implementation of all the methods mentioned in the form of user friendly web interfaces (www.metaboanalyst.ca) (Xia & Wishart, 2011a; Xia & Wishart, 2011b).

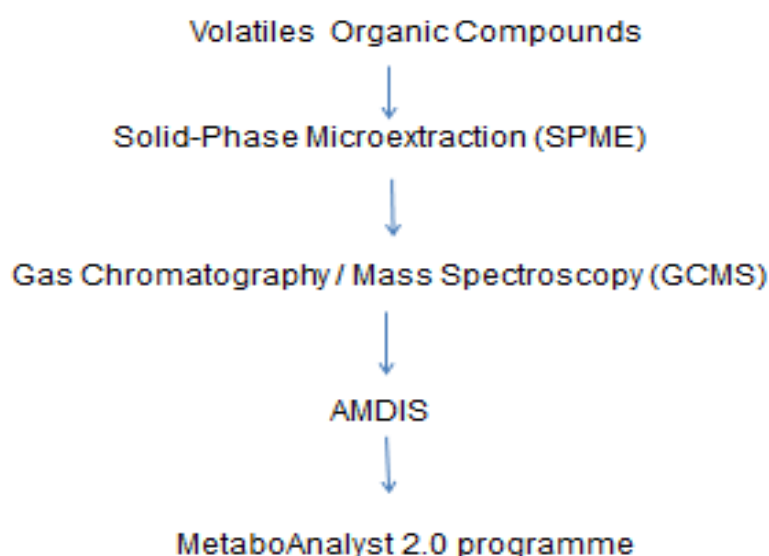


Figure 5.1 The flowchart of analysis of volatile organic compounds for identification of potential biomarkers of Cryptosporidiosis

5.3.6 Partial Least Squares - Discriminant Analysis (PLS-DA)

Partial least squares regression (PLS) is an extension of the multiple linear regression model. It is used to specify a relationship between a variable (Y), and a set of predictor variables, the X's, so that linear relationships may be analysed.

In order to see whether class discrimination was significant, I performed a permutation test. For each permutation, PLS-DA models were constructed using the data (X) and the permuted class labels (Y). The optimal number of components was used after determination by cross validation. MetaboAnalyst supports two types of test statistics for measuring the class discrimination. The first one is based on prediction accuracy during training. The second one (separation distance) is based on the ratio between the group sum of the squares and the within group sum of squares (B/W ratio). Class discrimination would not be considered significant statistically if the observed statistic is part of the distribution based on the permuted class assignments.

PLS-DA has two variable importance measures. 1) Variable Importance in Projection (VIP) the weighted sum of squares of the PLS loadings taking into account the amount of explained Y-variation in each dimension. 2) The weighted sum of PLS-regression. The weights are a function of the reduction of the sums of squares across the number of PLS components. When presenting more than two groups for analysis, each group will have the same number of predictors built. Therefore each features coefficient will be different. The importance of the overall coefficient is based on using the average of the feature coefficients.

5.3.7 Statistical analysis

Chi-Square was used to compare the variances of the sample group and control group in this study. The ratio of samples from each group of patients, that showed a characteristic were compared with the remaining proportion of those exhibiting the same characteristics. Additionally, MetaboAnalyst was used for data analysis. The presence or absence of each of the compounds found in a sample group and a control group, each stool sample was recorded “1” to indicate the presence of the data compiled and “2” to indicate the lack of each compound in that stool sample. The presence/absence data were used as the independent variables in a multivariate discriminant analysis with group membership.

5.4 RESULTS

A total of 118 volatiles were identified from 22 samples in the *Cryptosporidium* positive group (n = 11) and control group (n = 11). In all samples from both groups, five VOCs were shared by more than 80% of the subjects. The five VOCs common to more than 80% of the subjects were acetic acid and hexanal which were both found in 95.4% of samples (21/22). Ethylalcohol was found in 90.9% of samples (20/22), 2,3-butanedione in 90.9% of samples (20/22) and isopropyl alcohol in 86.4% of samples (19/22). The most predominant were acetic acid and hexanal (95.4%, 21/22), followed by ethylalcohol and 2,3-butanedione (90.9%, 20/22) and isopropyl alcohol (86.4%, 19/22) respectively. The frequency of each compound and the percentage are shown in Table 5.2.

5.4.1 Non-infected *Cryptosporidium* subjects (negative control group)

Analyses of the volatiles emitted from the stool specimens of the negative control group (n=11) demonstrated that the predominant VOCs were 1-propanol, acetic acid and hexanal. These three compounds were detected in 100% of subjects (11/11). Two other common compounds were 2-methylpropanal, and 2,3-butanedione which were detected in 90.90% of subjects (10/11). Moreover, other two compounds were ethylalcohol and isopropyl alcohol which were detected in 81.80% of subjects (9/11). However, only one compound was 1-propanol which was significantly different (100%, 11/11, Fisher's exact test, $p = 0.00$) when compared with negative control group (Table 5.2).

5.4.2 Specific changes in VOCs in *Cryptosporidium* positive samples

Nine compounds were shared between more than 80% of subjects: ethylalcohol (100%, 11/11), 2,3-butanedione (90.90%, 10/11), acetic acid (90.90%, 10/11), 3-hydroxy-2-butanone (90.90%, 10/11), hexanal (90.90%, 10/11), butanoic acid, 3-methyl (90.90%, 10/11), nonanal (90.90%, 10/11), and 2-pentylfuran (81.80%, 9/11). However, ethylalcohol found to be predominant of compounds in the *Cryptosporidium* positive group. Two compounds were 3-hydroxy-2-butanone (90.90%, Fisher's exact test, $p = 0.03$) and cyclopentane which were significantly different (63.6%, Fisher's exact test, $p = 0.01$) when compared

with both normal and sample group (Table 5.2). In *Cryptosporidium* positive samples, 86 volatiles were identified, with some inter and intra subject variations in VOC concentrations with time.

5.4.3 Statistical analysis of *Cryptosporidium* positive samples and negative controls

5.4.3.1 Analysis by Chi-Square test

Chi-Square was used to compare the number of each compound found in the *Cryptosporidium* positive samples to the compounds found in the negative controls (proportion). From the chromatograms produced by GCMS, 118 compounds in total between *Cryptosporidium* positive samples and the controls were identified (Figure 5.2). A Chi-Square test was undertaken and p-values were calculated for each compound (Table 5.2). From these results, 26 biological volatile compounds were determined from the highest frequency results both in *Cryptosporidium* positive samples and negative controls (Table 5.2). When compared to negative control group (88 volatiles were identified), volatile patterns from faeces of patients with *Cryptosporidium* samples were each significantly different. The fecal VOCs profiles of *Cryptosporidium* patients (n = 11) were compared with negative controls (n = 11). There were 2 VOCs (3-hydroxy-2-butanone and cyclopentane) which were significantly (Fisher's exact test, $p < 0.05$) more abundant in *Cryptosporidium* infected patients. There was 1 VOC (1-propanol) which significantly (Fisher's exact test, $p = 0.00$) more abundant in negative controls.

5.4.3.2 Partial Least Squares-Discriminant Analysis (PLS-DA)

Figure 5.3 demonstrates features identified by PLS-DA. It indicates the relative concentrations of the corresponding metabolite in the compounds found in the *Cryptosporidium* samples (group 1) and in the compounds found in the controls (group 2). It shows that a sample group has a high relation with cyclopentane and 3-hydroxy-2-butanone while it has no association with 1-propanol. In opposite, a control group has a strong relation with 1-propanol whereas it has no association with cyclopentane and 3-hydroxy-2-butanone. However, it is a very small sample size. Therefore, it might be a chance (95% probability, 5% chance).

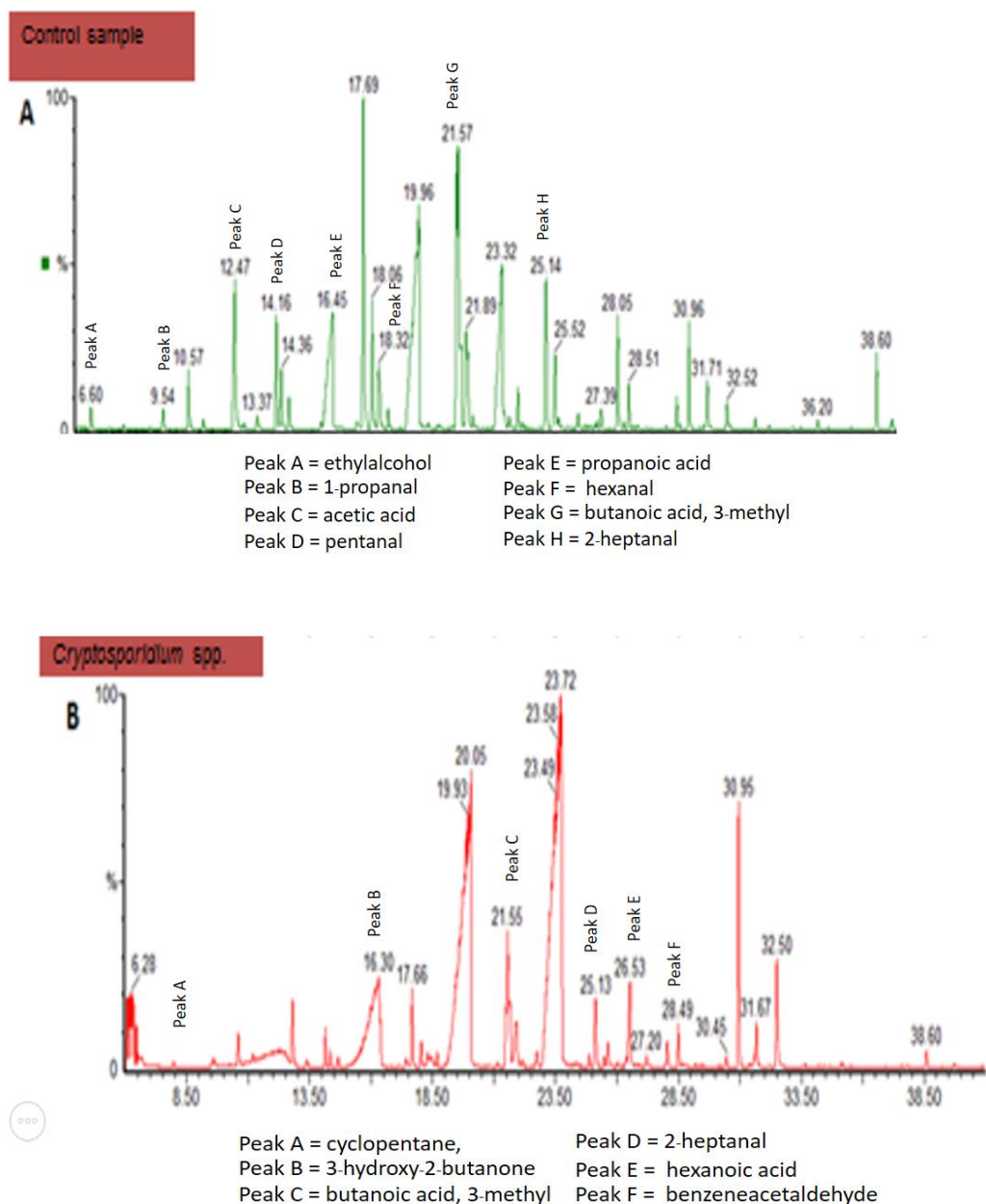


Figure 5. 2 Chromatograms of control sample and *Cryptosporidium* positive sample.

The figure shows two chromatograms produced by the GCMS analysis. Chromatogram A shows the results of a control sample (negative control, no *Crypto*) and chromatogram B shows the results for a *Cryptosporidium* sample (positive control). Both show percentage abundance of ions across retention time. Both possess a scanning mass range of 17 to 350 ions. Notice; an internal standard was not used.

Table 5.2 Biologically volatile organic compounds identified using AMDIS. Twenty six biologically compounds determined by Chi-Square test, these compounds have the highest frequency results either in *Cryptosporidium* samples or negative samples in order of retention times. Compound; 1-propanol was absent in *Cryptosporidium* samples but present in 11 control samples. Compounds; cyclopentane was present in 7 *Cryptosporidium* samples and only one in control samples and 3-hydroxy-2-butanone was present in 10 *Cryptosporidium* samples and 5 in control samples, these compounds have p-values of less than 0.05

No.	Compound name	RT	Freq. in crypto	Percent (%)	Freq. in Neg	Percent (%)	p-value
1	Ethylalcohol	6.6	11	100	9	81.8	0.24
2	Propanal	7.2	8	72.7	7	63.6	0.50
3	Acetone	7.4	6	54.5	7	63.6	0.50
4	isopropyl alcohol	7.6	10	90.9	9	81.8	0.50
5	Cyclopentane	8.3	7	63.6	1	9.10	0.01
6	2-methylpropanal	9.0	6	54.5	10	90.9	0.07
7	1-propanal	9.6	0	0	11	100	0.00
8	2,3-butanedione	10.3	10	90.9	10	90.9	0.76
9	acetic acid	12.6	10	90.9	11	100	0.50
5	butanal, 2-methyl	12.9	6	54.5	6	54.5	0.67
11	Pentanal	14.2	7	63.6	8	72.7	0.50
12	3-hydroxy-2-butanone	16.1	10	90.9	5	45.5	0.03
13	propanoic acid	16.3	5	45.5	6	54.5	0.50
14	Hexanal	18.3	10	90.9	11	100	0.50
15	butanoic acid	19.4	6	54.5	7	63.6	0.50
16	butanoic acid, 3-methyl	21.5	10	90.9	8	72.7	0.29
17	butanoic acid, 2-methyl	21.8	7	63.6	7	63.6	0.67
18	Heptanal	22.3	8	72.7	8	72.7	0.68
19	2-pentylfuran	24.8	9	81.8	7	63.6	0.32
20	2-heptanal	25.0	7	63.6	3	27.3	0.10
21	Benzaldehyde	25.6	8	72.7	5	45.5	0.19
22	2,2,4,4-tetramethyloctane	25.8	7	63.6	5	45.5	0.34
23	Octanal	26.1	7	63.6	6	54.5	0.50
24	hexanoic acid	26.4	4	36.4	6	54.5	0.34
25	benzeneacetaldehyde	28.7	6	54.5	5	45.5	0.50
26	Nonanal	29.5	10	90.9	6	54.5	0.07

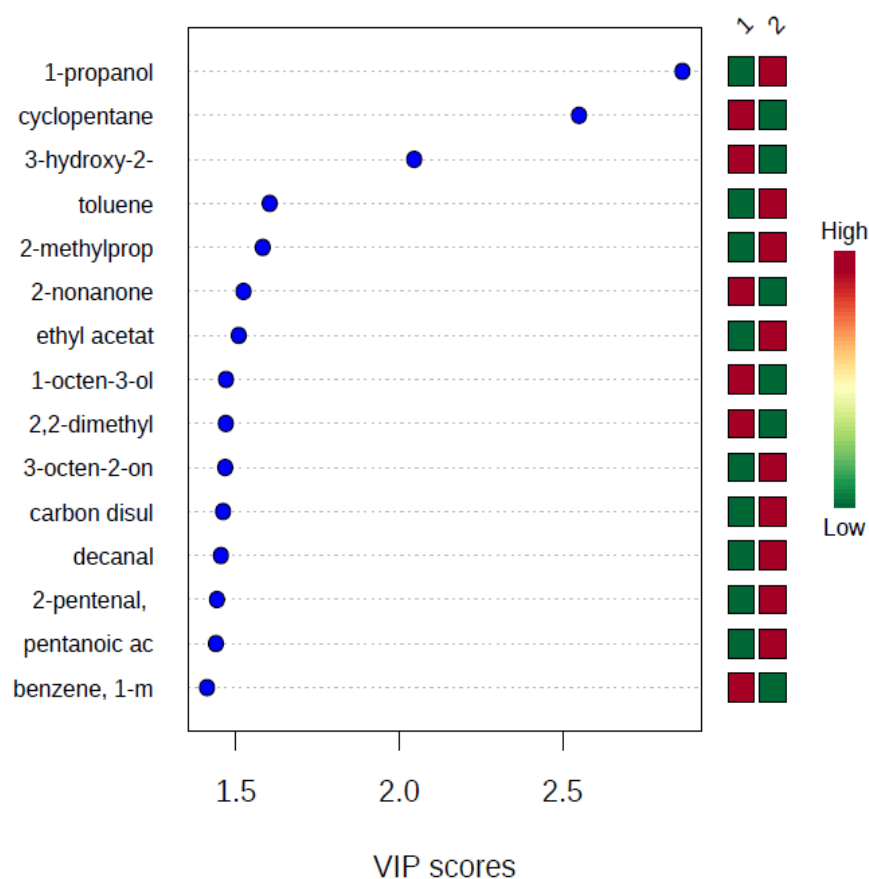


Figure 5.3 Shows the important features identified by PLS-DA. Variable Importance in Projection (VIP) this is the weighted sum of squares of the PLS loadings taking into account the amount of explained Y-variation in each dimension. It indicates the relative concentrations of the corresponding metabolite in the compounds found in the *Cryptosporidium* positive samples (group 1) and in the compounds found in the controls (group 2). Please note full names below from key used in graph;

3-hydroxy-2- = 3-hydroxy-2-butanone

carbon disul = carbon disulphide

2-methylprop = 2-methylpropanal

2-pentenal, = 2-pentenal(E)

ethyl acetat = ethyl acetate

pentanoic ac = pentanoic acid .benzene

2,2-dimethyl = 2,2-dimethyl-propanoic acid

benzene, 1-m = 1-methyl-2-(1-methylethyl

3-octen-2-on = 3-octen-2-one

5.5 DISCUSSION

In this study, I investigated the use of gas chromatography and mass spectrometry (GCMS) as a novel method for studying the VOCs in faeces of patients from Malawi with *Cryptosporidium* spp. infection. Therefore, I analysed and identified the volatiles emitted from the faeces of negative controls and patients with cryptosporidiosis and tested the hypothesis that VOCs are specific to the cause of diarrhoea.

The discovery of 2 biomarkers (cyclopentane and 3-hydroxy-2-butanone) for *Cryptosporidium* infection was a significant find in this area of volatile analysis. This technology by using “smell or odour” is able to differentiate the diarrhoeal diseases. The fact that these 2 compounds were then observed in the *Cryptosporidium* samples led to further analysis of the VOCs to try to distinguish between the two samples sets. 1-propanol was recorded to be absent in the *Cryptosporidium* samples but it was present in all negative samples. Therefore, it was concluded that this would show the difference between the two group samples. A clustering of cases drawn from the same population is clearly evident, and group to group differences are also shown.

This study found that cyclopentane and 3-hydroxy-2-butanone were present in increased abundance in the faeces of *Cryptosporidium* patients compared with negative controls. While the most commonly observed acid groups were acetic acid and hexanal which were more prevalent in both groups (95.45%, 21/22). The presence of VOCs such as ethylalcohol and 2,3-butanedione and isopropyl alcohol agrees with previous research (Parrett and Edwards (1997).

This result corresponds to the analysis of VOCs from the stools of neonates who were admitted at the NICU, Birmingham Women's, Hospital, United Kingdom (De Lacey Costello *et al.*, 2008). Three compounds, ethanol, hexanal and 2,3-butanedione were found with the highest frequency in both *Cryptosporidium* patients and negative control groups. However, three other compounds, acetone, 2-ethyl-1-hexanol and 3-methylbutanal were found in the UK samples whereas those compounds were not found in this study. Otherwise, for the other three compounds, 1-propanol, acetic acid and 2-methylpropanal were found only in this study, of the samples from Malawi.

From the results, we found that cyclopentane and 3-hydroxy-2-butanone were associated with *Cryptosporidium* positive samples, whereas 1-propanol was associated with negative controls. For example, in one case we found VOCs with high 1-propanol+low cyclopentane+low 3-hydroxy-2-butanone; this sample was negative for *Cryptosporidium*.

Previous studies described the identification of VOCs as biomarkers for rapid diagnosis of gastrointestinal diseases. Probert *et al.* (2004) investigated VOCs in stools from patients who had diarrhea (38 people) and healthy donors (6 people) by using GCMS as a novel method for a rapid diagnosis. Furans were the most prevalent of compounds in patients with *Clostridium difficile*, ethyl dodecanoate was the predominant one in patients with Rotavirus, ammonia was found to be ubiquitous in the VOCs from patients who had diarrhea with Norovirus and other enteric infections. However, in patients with *Campylobacter jejuni*, phenols were found to be the most predominant of the VOCs. Garner *et al.* (2007) analysed stool samples from 30 patients who were free of pathogens, 297 VOCs were found amongst the patient samples. The most common VOCs were acids, alcohol and esters. 44 of the VOCs were shared amongst 80% of the samples. It was noted that in patients with *C. difficile*, *C. jejuni* and ulcerative colitis the pattern of the VOCs was different to those without. Therefore, it was recommended that these different patterns could be used as biomarkers in clinical study. Additionally, other research by Garner *et al.* (2009) found VOCs from stools of Bangladeshi patients infected by cholera where fewer VOCs were identified in cholera samples in contrast to healthy volunteers. Dimethyl disulphide and p-menth-1-en-8-ol were the most prevalent from cholera specimens (Garner *et al.*, 2009).

Sharing of VOCs by negative controls, the VOCS emitted from the stools of asymptomatic donors consisted of 118 VOCs. Noticeably, more than 80% of donors shared 12 VOCs in the normal subjects. This finding confirms our hypothesis that several of the examined VOCs are shared. Several VOCs can be produced by the metabolism of micro-organisms in the small intestine or from foodstuffs. Other VOCs are remaining compounds from plants. A small number of VOCs are non-biological “pollutants” and reflect contamination of the environment (air, water, and food) by chemical waste. Interestingly, 5 compounds were shared between more than 80% of subjects in both *Cryptosporidium* sample group and the control group (acetic acid, hexanal, ethylalcohol, 2,3-butanedione and isopropyl

alcohol). There have been few reports of esters correlated with stool, which of course, does not possess the typical pleasant smell of esters. In the cohort study, 46 different esters were determined.

Several compounds that were present in the faeces of negative controls were also demonstrated in patients who had gastrointestinal disease. These shared compounds likely reflect the ongoing interaction between dietary substrates and intestinal flora. Nonetheless, distinct differences were found. The range of VOCs from stools of donors with ulcerative colitis, *C. jejuni*, and *C. difficile* infections was considerably less than for the total number of different VOCs of the stools from healthy volunteers. The abundance of *C. jejuni* and *C. difficile* organisms in the gut, with presumably relatively lower abundance of other microbial flora, might be expected.

This is the first time that the volatile organic metabolites in headspace gases emitted from the faeces of *Cryptosporidium* patients and negative control individuals have been analysed from samples collected in Malawi. Clear differences in the VOCs patterns were found that could differentiate between *Cryptosporidium* patients and negative controls. Having noted the microaerophilic nature of *Cryptosporidium* samples, it would be an interesting study to assess the hypothesis that due to 60% of the population having methanogenic bacteria in their intestine and that the ability for these bacteria to ‘use’ the methane requires oxygen, are these the same group of people who are more susceptible to *Cryptosporidium* samples. However, the biomarkers seen could lead to a rapid diagnostic device being made available for the analysis of the disease. The analysis of the volatile compounds that have been reported in this chapter are assisting in the understanding of the effects that infectious bacteria have on the human gut.

This study confirms our hypothesis, that the VOCs present in faeces acquired from patients suffering Cryptosporidiosis compared with faeces from patients without any pathogens will be different. This was shown initially by observing the difference in chromatograms produced by GCMS between *Cryptosporidium* positive and *Cryptosporidium* negative patient faecal samples. However, this was a very small sample size; a biomarker for *Cryptosporidium* samples was obtained from this small data set but maybe discounted in a larger study. In terms of the patterns of VOCs that are correlated with specific disease, more samples should have been collected (a larger group) from the same area in Malawi or in other countries to give a more representative study. This would lead to a more comprehensive analyses of the VOCs and a greater understanding of this disease process. This

technique of looking at the VOCs for disease diagnosis could contribute to the better understanding of disease progression and benefit the patient. Therefore, a larger scale study and/or a double blind study is necessary to confirm that these compounds are biomarkers. This is because the sample size may affect consistency and reliability of the results.

Comparison of the number of compounds detected in the control samples from Bangladesh showed only 87 VOCs (n=3) compared with the UK longitudinal study (n=30) this showed 298 VOCs whilst the number of this study was 11 samples, there were 118 VOCs. This may be due to the low numbers of samples from Bangladesh and Malawi. The number of compounds occurring at 100% in both sets of samples was considered. The samples sets from the Bangladesh and Malawi were so small when compared with a larger group of UK asymptomatic controls may be unsuitable. With such a small number of samples collected it is difficult to provide meaningful statistical analysis, a larger number of samples would be able to provide a more realistic analysis. More samples need to be obtained from Malawi or other countries to complete a comprehensive study into the patterns of VOCs that would lead to a broad analysis of the VOCs from these diseases. The controls from Malawi and from the UK were also compared; this demonstrated that there were similar compounds that occurred from both data sets although the reservation here is the very small data set for the Malawi controls. GC is generally accepted to be the best method for separating volatile compounds and is used throughout the world. It has several advantages over other methods, these include a rapid separation/analysis, it is relatively simple to use which keeps costs low, and the technique gives good, quantifiable results.

GCMS could potentially be used as a diagnostic test in Malawi and Cambodia as this method has the following advantages; it is fast and easy to perform in the laboratory (few steps are required); it can analyse lots of samples at the same time or in a short space of time (high throughput); and it is a reliable, tried and tested method.

In contrast, GCMS also has some disadvantages such as the following; it requires expensive equipment as it needs costly initial investment and highly technological equipment; the approach requires well-trained personnel; interpreting the results can take a lot of time as there are many procedures involved to ensure data is analysed correctly; and a large volume of faecal specimens is required.

However, if we can solve some problems concerning the disadvantages of using GCMS, I would recommend researchers to consider using this approach as a new and rapid diagnostic for causes of diarrhoea in Malawi and Cambodia. Particularly where there may be a delay of several days between the collection of a stool sample and a parasitological diagnosis. The delay may occur because of shipping of the samples to an appropriate laboratory but mainly due to the time required to complete the commonly used diagnostic techniques. The limited availability of expensive laboratory facilities capable of practising the parasitological, immunological and/or molecular techniques in developing countries further exacerbates the problem of establishing a diagnosis.

The assays could be modified to make them more suitable for use in developing country settings. For example, changing the SPME fibre and the protocol for GCMS could uncover further relevant volatile markers which could be even more sensitive and specific; the rapidity of the analysis of the VOC's is an area which I also believe could be greatly improved with further investigation. I have shown that faecal samples from patients with infectious diarrhoea contain characteristic chemicals that may be used to make a diagnosis without having to use laboratory techniques such as culture, ELISA, molecular techniques and/or electron microscopy. The use of gas from a simple headspace device should reduce the risk of cross infection in the laboratory. The development of an inexpensive device capable of near patient testing will hasten diagnosis, saving time and money. It is also possible that these techniques could also be applied to non-infectious gastrointestinal disorders (Probert *et al.*, 2004).

In conclusion, from the combination between developed advances in SPME with the novel method of the VOCs profile analyses from faecal specimens to examine the cause of infectious diarrhea, we have found changes in specific diseases, cyclopentane and 3-hydroxy-2-butanone were the most prevalent of the VOCs from patients with *Cryptosporidium* samples infection but 1-propanol was not associated with this group. In contrast, 1-propanol has a strong relation with normal controls. The VOCs found in faecal samples of *Cryptosporidium* patients compared to the VOCs present in pathogen free faecal samples shows them to be of significant difference ($p < 0.05$), showing potential for them to be biomarkers of the disease. These findings could lead the way to the development of a rapid diagnostic device based on VOC detection.

CHAPTER SIX

DISCUSSION AND CONCLUSIONS

In this study, molecular epidemiology was applied to *Giardia* and *Cryptosporidium* isolates identified in faecal specimens collected from children with diarrhoea in Malawi and Cambodia. My aim was to determine the range of *G. duodenalis* assemblage and sub-assemblage and zoonotic genotypes. Moreover, I investigated the prevalence and distribution of *Cryptosporidium* species in Malawi. Finally, I identified Volatile Organic Compounds (VOCs) in faecal specimens of Malawian patients who had *Cryptosporidium* infection.

In Chapter 2, I determined the prevalence of *Giardia* and *Cryptosporidium* in children in Malawi. I then went on to characterise the range of *G. duodenalis* genotypes, and identify the diversity of *Cryptosporidium* species, among children in Malawi. The overall prevalence of *G. duodenalis* in Malawi by 18S SSU-rRNA and *tpi* (at the same rate), *bg* and *gdh* were 28% (56/200), 10% (20/200) and 21% (42/200). The prevalence of *G. duodenalis* recorded from patient samples collected in Karonga, Lilongwe, and Mangochi by 18S SSU-rRNA and *tpi* were 23.4%, 30.6% and 29.7% respectively. Typing by the *tpi* gene, 100% (19/19) belonged to the sub-assemblage AII from all assemblage A. Among Assemblage B parasites, 87% (26/30) belonged to sub-assemblages BIII, 7% (2/30) displayed sub-assemblage BIV, 2% (1/30) belonged to sub-assemblage B (VB906855) and 2% (1/30) showed sub-assemblage other B (unidentified). The prevalence of *G. duodenalis* recorded from patient samples collected in Karonga, Lilongwe and Mangochi by using the *bg* gene, was 17.2%, 9.7% and 3.1%, respectively. Typing by the *bg* gene, 33% (1/3) contained sub-assemblage AI, 67% (2/3) displayed sub-assemblage AII. For Assemblage B, 46% (5/11) contained sub-assemblage B heterogenous, 27% (3/11) showed sub-assemblage BIII and 27% (3/11) belonged to sub-assemblage other B (unidentified). The prevalence of *G. duodenalis* recorded from patient samples collected in Karonga, Lilongwe and Mangochi by using the *gdh* was 21.9%, 20.8%, and 25.5%, respectively. Typing by the *gdh* gene, 56% (18/31) contained sub-assemblage BIV H43 (EF507682), 38% (12/31) displayed sub-assemblage BIV (KC96064.1) and 6% (2/31) belonged to sub-assemblage B Cla145 (HM134212.1).

The overall prevalence of *Cryptosporidium* spp. in Malawi by SSU-rRNA gene was 11% (22/200). The prevalence of *Cryptosporidium* spp. recorded from samples collected in Karonga, Lilongwe and Mangochi was 6.3%, 22.2%, and 3.1% respectively. The overall prevalence of *Cryptosporidium* spp. in Malawi by GP60 gene was 8% (16/200). The prevalence of *Cryptosporidium* spp., recorded from samples collected in Karonga, Lilongwe and Mangochi was 6.25%, 13.88% and 3.125% respectively. Restriction analysis of the SSU PCR products showed that 10 of the 22 PCR-positive samples had *C. hominis*, eight had *C. parvum* and four samples were infected with both species (*C. hominis*/*C. parvum*). Sequence analysis of the SSU rRNA and GP60 genes confirmed the species identification by SSU rRNA-based PCR-RFLP analysis. Mixed infection (*C. hominis*/*C. parvum*) was found in 25% of samples from Lilongwe. Moreover, *C. hominis* and *C. parvum* were found in the same percentage (37.5%). However, only *C. hominis* was found in Mangochi whereas both species were found in the same percentage (50%) in Karonga.

While there are a paucity of data describing *Giardia* and *Cryptosporidium* infections in Malawi, my study was limited in that I examined only specimens taken from children with diarrhoea. Since I was not able to access specimens from asymptomatic children, I was not able to determine the true incidence of Giardiasis and Cryptosporidiosis in the population.

There were published studies of both symptomatic and asymptomatic cases in African countries. Genotyping of *G. duodenalis* isolated from symptomatic and asymptomatic infected Egyptian school children by using Nested PCR targeting intergenic spacer (IGS) found that Sub-assemblage AI was the commonest with 66.6% (20/30) among asymptomatic children compared to 53.3% (16/30) of symptomatic children, while assemblage B was found in 40% (12/30) of symptomatic compared to 20% (6/30) of asymptomatic. The difference was significant. Assemblage AII was only found in asymptomatic with 13.4% (4/30), while mixed infections (AI&B) were recorded only in 6.6% (2/30) of symptomatic group. A significant relationship was found between younger children susceptibility for AI and B infections as presented in 77.7 (12/16) and 83.3% (10/12) of symptomatic children, respectively, and 80 (16/80) and 33.4% (2/4) of asymptomatic children, respectively. Significant relations were found between AI with intermittent diarrhea and B with chronic. A significant relationship was found between assemblage distributions and heavy infection intensity.

In conclusion, higher incidences of assemblage B among symptomatic children compared to asymptomatic children could denote its possible pathogenic potential. (Hussein *et al.*, 2016)

Cryptosporidium infection rates amongst children are similar to those found during other studies in sub-Saharan Africa, including the Global Enteric Multicenter Study (GEMS) and in the study of Mbae *et al.* in Kenya (Kotlof *et al.*, 2013, Mbae *et al.*, 2013). The higher prevalence in cases than in controls agrees with findings from Kenya and supports the idea that the parasite causes symptomatic diarrhoea (Mbae *et al.*, 2013). However, the study by Vargas *et al.* found that only one out of 451 hospitalized children < 5 years of age in Kilombero district in Tanzania had infection with *Cryptosporidium* (Vargas *et al.*, 2004). This could be due to differences in methodology, using microscopy which is generally known to be less sensitive, although regional differences may exist. Cryptosporidiosis can often have severe effects, on those whose immune systems are compromised such as HIV+ patients. Many studies have focused on its prevalence amongst such patients usually adults but also sometimes including children. Among study participants with known HIV-status, including both cases and controls, *Cryptosporidium* infection was significantly more prevalent in HIV-positive (24.2%, 8/33) than in HIV-negative (3.9%, 15/387) children in univariate analysis ($p < 0.001$; OR=7.9; 95%CI:3.1–20.5). In West Africa, however, cryptosporidiosis is the cause of diarrhoea in 4.9% to 14.7% of immunocompetent children, although this can vary according to age and area (Sulaiman *et al.*, 2003). Similarly, our results showed that *Cryptosporidium* infections were strongly associated with diarrhoea in children under five years of age and was present in 11% of symptomatic children at the age less than 1 year. Asymptomatic *cryptosporidium* carriers were not observed in this study. A review of cryptosporidiosis in sub-Saharan Africa reported the same age distribution with a peak amongst children aged 6–12 months. Infection can occur throughout childhood but symptoms lessen with age (Verweij *et al.*, 2004).

The majority of *Cryptosporidium* positive samples were from children younger than one year, but age was not a significant factor. This might not be very surprising since all the study participants were below 2 years of age, and other studies which also included older children have reported a higher prevalence of infection among younger children, those below 2 years of age in particular (Gatei *et al.*, 2006; Kotlof *et al.*, 2013; Mbae *et al.*, 2013).

PCR-RFLP analysis showed at least four species of *Cryptosporidium* in Kenya; *C. hominis*, *C. parvum*, *C. felis* and *C. meleagridis*. *C. hominis* which is almost exclusively a human parasite (Xiao, 2010) was the most common with a prevalence of 82.7%. The other three are zoonotic species but are often associated with human cryptosporidiosis (Xiao, 2010). Previous studies on the prevalence of *Cryptosporidium* species and genotypes infecting children in Kenya also found that *C. hominis* was the dominant species. Moreover, *C. parvum*, *C. meleagridis*, and *C. muris* was detected in HIV-infected persons (Gatei *et al.*, 2002; Gatei *et al.*, 2006). *C. parvum* was shown to be more of a severe disease than *C. hominis* as it can be found more in patients (63%) than outpatients (37%) while *C. hominis* were found in (54%) of outpatients compared to inpatients (46%). Cama *et al.*, (2007) also reported that *C. parvum* was more pathogenic than *C. hominis*. The presence of zoonotic species including *C. felis*, and *C. meleagridis* would indicate that animal reservoirs are still important. The distribution of *Cryptosporidium* species and genotypes in a population is a good indicator towards sources of potential infection, therefore contact between humans will probably have played an important role in the distribution of *Cryptosporidium* in this study. Conversely, the mode of transmission cannot be definitely concluded whether passed from animal to child or via contact with contaminated food, water or with animal faeces. The results of this study back up previous work amongst the child populations of Africa and the developing countries. These studies show that 79 - 90% of infections are caused by *C. hominis* (Ajjampur *et al.*, 2010; Molloy *et al.*, 2010; Xiao and Feng, 2008). In Malaysia and Kuwait, the same four species were identified, however in their study, *C. parvum* was the most frequently detected species followed by *C. hominis* (Lim *et al.*, 2011). *C. felis* is one of the five most common *Cryptosporidium* spp. that are responsible for human cryptosporidiosis. *C. felis* was first recorded among humans in Kenya by Gatei *et al.*, (2006b). *C. meleagridis*, one of the zoonotic species identified in this study has been shown to be a human pathogen in Africa (Kenya), Peru, India and Thailand (Xiao *et al.*, 2001; Gatei *et al.*, 2006b). *C. muris* and *C. canis* identified in an earlier study by Gatei *et al.* (2006b) were not among the species found in this study. However, possible asymptomatic *C. muris* infection in healthy persons has been (Sharma *et al.*, 2013) and in an immunocompromised patient (Gatei *et al.*, 2006b) leads to the suggestion that this maybe another *Cryptosporidium* species with zoonotic potential.

For genotyping of *Cryptosporidium*, Moore *et al.* (2016) reported *Cryptosporidium* species identification in Cambodia, six *Cryptosporidium* species were detected. Of the 38 *Cryptosporidium* positive samples, 13 (34%) were *C. hominis*, 9 (23.7%) were *C. meleagridis*; 8 (21.1%) were *C. parvum*, 5 (13.2%) were *C. canis*, and single positives found for *C. suis* and *C. ubiquitum*. One sample was positive with a mixture of *C. hominis* and *C. parvum*. Six different *Cryptosporidium* species were found in Cambodian children. The most common was *C. hominis*, a common, anthroponotic species in developing countries (Xiao, 2010). The researchers suggested that anthroponotic transmission of *C. parvum* is important in Cambodia. Moreover, they also found nine patients with *C. meleagridis*, five patients with *C. canis* and one patient each with *C. suis* and *C. ubiquitum*. These species are also found in animals so while anthroponotic transmission of *Cryptosporidium* appears to be most important in children in Cambodia, it also points to there being a zoonotic reservoir of infection.

With *C. parvum* and *C. hominis* infections in the paediatric population studied there, further examination of the genetic variation within each of the two species was undertaken using partial gp40/15 or GP60 subtyping tool, which at the moment is the most common locus for identifying *Cryptosporidium* subtype families and subtypes. Knowledge of the subtypes of *C. hominis* and *C. parvum* may help understanding into the modes of transmission and infection of these organisms and help strategies for effective prevention and treatment.

Molecular characterisation of *Giardia* and *Cryptosporidium* species in humans has been undertaken using PCR-based typing techniques. From the past two decades, SSU-rRNA has been extensively used to classify *Giardia* and *Cryptosporidium* spp. as well as being used for the characterisation of *G. duodenalis* assemblages and genotyping of *Cryptosporidium* spp. However, using only one locus or two loci (SSU-rRNA, *bg*, *gdh* or *tpi*) or one marker was not totally successful in all specimens. Moreover, some specimens were hard or problematic to amplify and/or provided unsatisfactory results (bands of PCR products were faded). This led to ineffective sequencing. The conditions of the PCR reaction and the quality of the DNA specimens always affected the outcomes. Where ineffective sequencing and poor results occurred it was necessary to perform optimization of the PCR to improve the outcome. This is where the reaction can be modified by changing the non-constants such as the concentration of the reagents and states of the cycling (Lorenz, 2012). Varying the stringency of the reactions somewhat can lead to differing results, too much stringency can generate many false

amplicons, and too little may lead to none (Lorenz, 2012). Amplification may also be problematic where denatured DNA or structures formed with the primers are unable to bond with the correct DNA counterpart. Furthermore, genetic variation of *Giardia* and *Cryptosporidium* spp. might influence the result of the PCR technique.

In overall conclusion, this study revealed information about the distribution of *Giardia* assemblages/sub-assemblages and *Cryptosporidium* genotypes in humans. However, it was not representative of the true prevalence of Giardiasis and Cryptosporidiosis. This is because all samples were collected from symptomatic cases only. Thus, may reveal a higher infection rate than a study of normal samples. To better understand the genetic diversity of these parasites, a combination of markers is required. Prospective work should undertake collecting a larger group of samples from both symptomatic patients and asymptomatic cases. Also the development and use of new and a variety of markers such as microsatellites would help to more comprehensively describe parasite diversity. In addition, fresh samples would help to avoid problems from DNA degradation, such as very low DNA extraction and subsequent low yield of PCR products. I also attempted to sequence the genome of *Cryptosporidium* spp. using DNA extracted from oocysts that were purified from faeces via immunomagnetic separation. However, this method did not succeed as we could not see any oocysts, possibly due to the presence of substances that inhibited DNA extraction.

In Chapter 3, I investigated the diversity of *G. duodenalis* in various regions of Cambodia using multilocus genotyping. By using this technique targeted at various genes, a single marker was used for characterisation of the *G. duodenalis* assemblages, and particular markers such as the *bg*, the *gdh* and the *tpi* gene was used for subtyping at the sub-assemblage level. Overall, amplification successes differed according to the locus examined. The *tpi* assemblage-specific gave the highest percentages of the results whereas the *gdh* and *bg* loci gave lower percentages of the results. *Giardia* is a complex species as it has been categorised into at least eight (A to H) morphologically identical genotypic assemblages with variable host occurrence. Therefore, using various markers is an appropriate way to study the genetic variation of *G. duodenalis*.

This is because *Giardia* has been shown to be a complex species. In this study, the *tpi*, *bg* and *gdh* genes were selected to analyse human samples. Using the multilocus genotyping approach, the

current study identified sub-assemblages of *G. duodenalis* in humans. A high level of polymorphism was detected, confirming that *G. duodenalis* is indeed a complex species. Some isolates could not be classified to the exact sub-assemblages, which could suggest that these groups are novel sub-genotypes, particularly assemblage B in humans.

I performed for the first time a multi-locus genotyping approach to identify the *Giardia* assemblages and sub-assemblages in patients from Cambodia. Using the standard multi-locus sequence typing approach based on the *bg*, *gdh* and *tpi* loci, this study classified sub-assemblages of *G. duodenalis*. This result agreed with many previous reports that Assemblage B had a strong correlation with clinical symptoms of *Giardia* infection as it was more common in symptomatic patients than asymptomatic patients (Volotao *et al.*, 2007; Mohammed Mahdy *et al.*, 2009; Sprong *et al.*, 2009;). Conversely, this study disagreed with a few reports that found a strong association with assemblage A with clinical symptoms (Haque *et al.*, 2005; Sahagun *et al.*, 2008; Perez Cordon *et al.*, 2008). The results correspond to a study from India, where the highest prevalence of *G. duodenalis* was 15.5% (74/484) by *tpi*, followed by *gdh* (10.54%) and *bg* (9.92%). The highest prevalence of the *G. duodenalis* assemblage was assemblage B in both adults and children (82.4%) while assemblage AII was only 9.4% and assemblage AI was not detected (Laishram *et al.*, 2009). Assemblage B parasites belonged to sub-assemblages BIII and BIV whereas all assemblage A belonged to the sub-assemblage AII (Laishram *et al.*, 2012; Sprong *et al.*, 2009). I also found a high level of polymorphism and sub-assemblages in this highly diverse species. It would be useful to collect stool specimens of humans and their animals from the same area, to provide valuable information about possible zoonotic transmissions.

To better understand the molecular genetics of *G. duodenalis*, a combination of techniques may be required for instance, MLG and proteomics analysis. For this study, genotyping by using *tpi*, *bg* and *gdh* markers was not completely successful in all samples. This may be due to the quality of DNA samples that affected the PCR results. Thus making the samples difficult to amplify. Moreover, the PCR products had a very low concentration, sequencing would not always be successfully. Otherwise, the genetic variation of *G. duodenalis* may also affect the outcome of the PCR products at a particular marker, producing a variety of results. For future work, the MLG of *G. duodenalis* could be performed and improved by adding more molecular markers, developing new markers and

new techniques and finally, the evaluation of *G. duodenalis* proteomics may also be necessary. Moreover, studying both cases and control is crucial, so for the most effective work, a large sample size of both normal cases and symptomatic samples is required.

In conclusion, the results of molecular epidemiology of *Giardia* in samples from Cambodia agreed with many previous studies, not only from Southeast Asia but also corresponding with those found in various regions around the world (Volotao *et al.*, 2007; Mohammed Mahdy *et al.*, 2009; Sprong *et al.*, 2009). Assemblage B was associated with the presence of clinical symptoms, followed by assemblage AII infection. However, *Giardia* parasites of assemblage B demonstrated higher genetic variation than the other assemblages (assemblages A and mixed infection). The main *G. duodenalis* subtype was assemblage B. Further work is needed to fully describe the epidemiology of disease and the circulating subtypes *G. duodenalis* in Cambodia. Moreover, the sources of zoonotic *Giardia* and routes of waterborne exposure both in the winter (dry) and in the rainy (wet) season in this region should be determined. To further clarify the source and routes of zoonotic transmission, future works should investigate the prevalence and species of *G. duodenalis* in wildlife and domestic animals and additional routes of waterborne exposure.

In Chapter 4, I applied for the first time a MLST technique to characterise the *Giardia* assemblages and sub-assemblages in patients from Malawi and Cambodia. I showed that the molecular diversity of *Giardia* assemblages in patients from Malawi and Cambodia corresponds to many studies observed worldwide (Sprong *et al.*, 2009) and in other parts of Southeast Asia (Volotao *et al.*, 2007; Mohammed Mahdy *et al.*, 2009; Sprong *et al.*, 2009). I confirmed that *Giardia* assemblage B is more prevalent than assemblage A in symptomatic disease, and I demonstrated the presence in our patients of multi-locus genotypes observed in other parts of Southeast Asia. Despite allowing a direct comparison of our results with those from other studies, the use of the standard MLST suffered from a series of limitations. These included the lack of amplification at particular loci in certain isolates and the high frequency of occurrence of sequences with heterogeneous positions. The accurate genotyping of assemblage B isolates in particular was problematic, and the need for developing more reliable molecular tools for the genotyping of parasites of this assemblage is evident. In this study, the maximum likelihood (ML) was performed in order to demonstrate the order of isolates based on the nucleotide homology of the sequences. To determine genetic variation among

assemblages and sub-assemblages, the sequences of the *tpi*, *bg* and *gdh* gene fragments from each isolate were merged, a multiple alignment were performed and phylogenetic trees were generated using ML method. I found the genetic diversity in some samples both in Malawi and Cambodia. Phylogenetic analysis of multilocus markers grouped the isolates into two distinct clades; clade I and II. Clade I which contained assemblage B, which the mainly of the isolates clustering within this clade. Clade I; was composed of two main sub-clades I-II. Sub-clade I which contained assemblage BIII with isolates from Malawi in 53.1% (26/49) and isolates from Cambodia in 30% (12/40) by the *tpi* gene. For the *bg* gene showed isolates from Malawi in 21.4% (3/14) and isolates from Cambodia in 31.6% (6/19) whereas were not found any isolates from both regions by *gdh* gene. Sub-clade II was composed of assemblage BIV with isolates from Malawi (8.2%, 4/49) and isolates from Cambodia (40%, 16/40) by the *tpi* gene. For the *bg* gene showed isolates from Malawi in 57.1% (8/14) and isolates from Cambodia in 47.4% (9/19) whereas the *gdh* gene demonstrated isolates from Malawi in 70% (28/40) and isolates from Cambodia in 53.5% (23/43). Clade II contained assemblage A with isolates from Malawi (Sub- assemblge AII in 38.8%, 19/49) and isolates from Cambodia (Sub- assemblage AII 27.5%, 11/40 and AI 2.5%, 1/40) by the *tpi* gene. For the *bg* gene showed isolates from Malawi in 57.1% (8/14) and isolates from Cambodia in 47.4% (9/19) whereas the *gdh* gene showed only isolates from Cambodia in 16.3% (Sub-assemblage AII 13.9% (6/43) and AI 2.3% (1/43). The *tpi* was more able to distinguish giardiasis, mixed assemblages and sub-assemblages than the *bg* and *gdh* loci. The molecular genotyping results showed that *Giardia* assemblage B was responsible for the majority of the clinical infections and confirmed the occurrence of a high diversity of parasite multi-locus genotypes.

The basis for the difference between isolates of *Giardia* from Malawi and Cambodia is unknown but again as in the case of human host factors such as immunity, behavior, susceptibility and genetic background. For environment factors such as culture, life style, traditional food, personal hygiene, socio-community and geographical condition. Moreover, parasite factors such as mutation, adaptation of themselves against anti-parasitic drugs and natural evolution depends on geographical location.

In most cases characterisation of the molecular diversity of assemblage A isolates was successful. Heterogeneous positions amongst sequences was only detected in a small number of isolates, mostly

at the *bg* locus. This collaborates the unusual occurrence of nucleotides overlapping as regularly reported in assemblage A parasites (Sprong *et al.*, 2009). Whichever gene was subjected to analysis, nearly all sequences were matches for previously described isolates. Novel polymorphisms were very rarely observed. Almost all assemblage A isolates assigned to sub-assemblage AII throughout the three loci. Confirmation of the far greater number of sub-assemblage AII as opposed to sub-assemblage AI in humans was established and this was supported by previous multi-locus sequence typing studies (Caccio *et al.*, 2008; Geurden *et al.*, 2009a; Lebbad *et al.*, 2011).

The differing prevalence of either assemblage A or assemblage B could be due to the fact that the two assemblages (A and B) vary in their levels of dominance from one area to another. Differing levels of susceptibility could also be accredited to differing levels of sanitation, leading to greater increased risk of infection. It may also be possible that effective immunity has been unable to develop. It remains to be seen why there is such a variation from region to region in the predominance of the *Giardia* assemblages. It could be related to the fact that the two assemblages tend to favour differing methods of transmission. Assemblage A more often transmits zoonotically whereas assemblage B is more often transmitted from human to human although it may also have the potential to transmit zoonotically.

There appears to be an even greater and complex number of assemblage B isolates in comparison to assemblage A and with an even greater number of sequences with heterogeneous positions that had overlapping nucleotides, these came out at about 20% as an average across the three loci. Sprong *et al.* (2009) noted how common it was for assemblage B parasites to have mixed templates and this was demonstrated here by some of the isolates being unable to be definitively resolved. Ankarklev *et al.* (2012) noted their presence following DNA extraction and sequencing even from one *Giardia* cyst from the same human. The assemblage B was the most predominant in this area and there no animal assemblages were found in these samples. However, it is important to collect samples from animals to investigate the true prevalence of zoonotic transmission.

Different markers showed discrepancies between some sub-assemblage B isolates meaning that they could not always be immediately identified. At the *bg* locus most isolates belong to subtypes from the previously recognised B1 group (Geurden *et al.*, 2009a), having been previously seen in human isolates from Sweden (Lebbad *et al.*, 2011). Verification of a B1 cluster near the sub-assemblage BIII

was observed by phylogenetic analysis but whether this is or represented another assemblage rather than a BIII or BIV could not be confirmed. Genotypic diversity was seen to be much higher in assemblage B isolates than assemblage A isolates. Assemblage B exhibited 21 unique sub-types most of which were at either the *gdh* or *tpi* loci. Studies of European patients by both (Broglia *et al.*, 2013; Lebbad *et al.*, 2011) also found levels of diversity that were similar in assemblage B parasites in humans. A number of infected patients presented with sub-assemblage AII multi-locus genotypes (AII-1 and AII-2). A number of studies have already identified these MLGs in patients expressing symptoms, these studies include in Italy (Caccio *et al.*, 2008), Belgium (Geurden *et al.*, 2009a) and Sweden (Lebbad *et al.*, 2011). All these studies have demonstrated that there is a large and diverse number of assemblage A genotypes that are able to infect humans.

After successful characterisation of all three markers following multi-locus analysis, infection amongst most patients was confirmed to be the same multi-locus genotype as previously observed by (Lebbad *et al.*, 2011) in Sweden. High numbers of assemblage B, MLGs were discovered demonstrating different combinations and never been seen before combinations of *bg*, *gdh* and *tpi*, similar to assemblage A. The current study showed the complex sub-assemblages of *G. duodenalis* isolates from Malawi and Cambodia based on *tpi*, *bg* and *gdh* loci.

In conclusion, *tpi* marker are a highly resolving and helpful tool for investigation of variability within the *Giardia* in the same area. They provide a simple technology which can be used to distinguish areas and this forms a foundation to explore further the cause of the genetic variation. This study opens up an arena for formal genetics, population genetics, evolution and epidemiology which should contribute to the understanding of this important parasite. The result of this study has for the first time demonstrated on existence of genetic variation among the *Giardia* isolates from Malawi and Cambodia. The cause of this observed genetic variation of *Giardia* is obviously multifactorial and thus this deserves further investigation for a better understanding on the biology of this particular parasite. It remains to be seen whether this observed level of diversity is correlated with any phenotypic characteristics or the virulence of the parasite in various endemic communities in Malawi and Cambodia.

In Chapter 5, I used a novel method (GCMS) to examine the VOCs from human faeces from Malawi that contained *Cryptosporidium*. My hypothesis that the pattern of VOCs in diarrhoeal stool may

differ by aetiology was based on the disease specific “odour” that is known to medical staff caring for patients with gastroenteritis. The number of VOCs studies in the context of protozoan parasites is small. The current study aimed to describe the pattern of VOCs emitted from faeces with the purpose of identifying biomarkers for pathogen- specific diseases.

The technique for extraction of VOCs has been developed especially for use in faecal samples. The VOCs have been catalogued from faeces in negative control (non-infected controls). Statistical analysis has been used to compare the VOCs of infectious diarrhoea by *Cryptosporidium* spp. from humans. The difference of VOCs will be correlated with various kinds of gastrointestinal disease status and the VOCs will have a great diagnostic value for examining disease. However, a small group of study in Malawi has been investigated. The extraction and concentration method SPME which was utilised in this study is different from other extraction methods. So, it may be possible that some VOCs emitted have not been extracted. Nevertheless, there were only a few published papers concerning VOCs. The advantage of SPME is the simplicity of the pre-concentration step.

GCMS is a proven analytical method for the analysis of VOCs so it was not examined in this thesis. The improved chromatography received from utilising a conjoined capillary column showed that the GCMS technique warranted some development in the analysis of faeces because of the difference of polarities of the VOCs in the sample. It allowed for better identification of the peaks in the chromatogram using NIST 05 and retention time matching. The major disadvantage of SPME with the CAR/PDMS fibre is the non-quantitative or semi-quantitative nature of the fibre. It may be difficult to compare this work with other quantitative studies. Nonetheless, the fact remains that the higher amounts of VOCs that have been examined would help in the development of this work. The catalogue of VOCs in non-infected controls, demonstrates that there are some VOCs shared with the normal stool samples. It supports the hypothesis that a basic number of VOCs could be used as a gold standard of health, in that diagnosis could be quicker, more cost effective and accurate. This means that the VOCs would be common in healthcare and that VOCs would prove to be constant amongst individuals.

The discovery of two biomarkers for *Cryptosporidium* spp. was a significant finding of my work, demonstrating the possibility for detecting disease states utilising the “smell technology”. This would

lead to an easy and quick diagnostic tool being available for the investigation of the disease. The compounds have the ability to be used statistically in order to determine four distinct groups. This could be employed for pattern recognition or sensor arrays to differentiate the specimens. A larger number of asymptomatic specimens should be analysed for different dietary groups. However, the number of samples was small and all people were omnivorous in this thesis so we could not compare the differences between various dietary groups or other clinical information. Further research should incorporate a larger and greater variety of groups with differing diets so that dietary bias differences can be avoided. The current study demonstrates the potential of VOCs in detection of *Cryptosporidium* spp. Furthermore, the principle based on the disease specific “smell” has been supported and it could be a very useful device to enable non-invasive detection.

Only *Giardia* assemblage A and B are known to be related to human infection. Different countries and different regions within those countries have been shown to have varying prevalences of each assemblage. It is unknown as to why the distribution of assemblages varies so much.

In the current study, DNA sequencing showed that mixed infections with both *Giardia* assemblage A and B were predominant when using the *tpi* gene. However, assemblage B was predominant when using the MLST loci. The predominance of *C. hominis* indicated that the anthroponotic route plays a major role of *Cryptosporidium* transmission in Malawi. Both assemblages A and B are common in Cambodia, however, assemblage B was the most predominant. It indicated that anthroponotic routes play a major role of *Giardia* transmission in Cambodia. Nevertheless, *Giardia* parasites of assemblage B demonstrated higher genetic variation than the other assemblages. The major transmission route is anthroponotic in *Giardia* (83.3%). This has potential implications for public health policies to lower the incidence of disease. If person-to-person transmission is dominant, improvements in sanitation and hygiene recommended, rather than changes in animal/meat processing.

CONCLUSIONS

Giardia Assemblage B was predominant in both Malawi and Cambodia. A number of factors could explain the prevalence of the infection rates, these include local/regional epidemics, social and economic factors and the methods of transmission due to local hosts and sources of infection.

The results of this study proposed anthroponotic transmission from human to human. However, no study into genotype assemblages A and B from animal samples was undertaken, therefore human and animal isolates weren't compared, this has been one limitation of the research. Therefore, the current results were not enough to demonstrate the role of anthroponotic and zoonotic transmission of giardiasis infections. Further study is required into human and animal populations that will identify the molecular genotypes and epidemiological information, this should demonstrate the distribution and prevalence of *G. duodenalis* assemblages and subtypes within human and animal populations. Work to identify prevalences of assemblage A or B in domestic and pet animals should be undertaken to help identify possible reservoirs of zoonotic giardiasis infection for humans. Also a thorough more molecular characterisation at multiple loci shall be performed in the future on *G. duodenalis* isolates ranges from humans, animals and the environment.

Several studies have shown these infections to be linked to socio-demographic factors including sanitation and nutritonal levels and host immunity levels and the parasitic strain (Thompson *et al.*, 2000; Hoque *et al.*, 2001). Giardiasis infections have also been associated with the movement of peoples to and from endemic areas and the consumption of both contaminated tap and fresh water (Hoque *et al.*, 2001). Antenatal studies in Auckland, New Zealand have shown that changing babies nappies was associated with a risk of *Giardia* infection (Hoque *et al.*, 2001). In developing countries, common risk factors associated with *G. duodenalis* infection are often characteristic of the population: poor levels of sanitation, inadequate cleansing of the person, consumption of unwashed produce and drinking of contaminated tap water (Stuart *et al.*, 2003; Mohammed *et al.*, 2003). Mukherjee *et al.* (2009) showed that in Kolkata, India, there was a relationship between giardiasis and the socio-economic background of the study population. The majority of the patients suffering diarrhoea were of a lower socio-economic status and lived in slums, suggesting that infection could

be water and/or food borne. Further, the disease appears to be more common amongst children, children less than 5 years of age were most at risk of parasitic infections. Rapid growth of urban areas without access to treated, clean water means that large numbers of people are left to drink potentially infected, untreated water. Areas such as day-care centres and schools are possible sources of contamination where large numbers of children and are in close contact (Atta-Owus *et al.*, 2008). Therefore, increased public health education focusing on good sanitary practices, particularly among mothers and children, could decrease the risk of *G. duodenalis* infection.

The current study is the first to describe the anthroponotic and heterogeneous nature of *Cryptosporidium* species infecting children from Malawi. *C. hominis* was the most common species with two different subtypes, followed by *C. parvum* with a single subtype. Further work is needed to fully describe the epidemiology of disease and the circulating subtypes of *Cryptosporidium* in Malawi and *G. duodenalis* in Cambodia. This current study involved children living in the urban informal settlements of Malawi and Cambodia, where living conditions are often overcrowded and the populations are poor and have limited access to decent sanitation and hygiene, however direct contact with animals is often limited or virtually non-existent. The distribution of *Giardia/Cryptosporidium* species and genotypes in a population is an indication of the potential infection source. Therefore, infection transmission from human to human is likely to have played an important role in Giardiasis/cryptosporidiosis epidemiology in the children in this study. It could suggest that the anthroponotic species of *Giardia/Cryptosporidium* were transmitted directly from child to child, via contamination of water, food, or hands with human faeces. These findings are consistent with results from previous studies in paediatric populations in Africa (Ajjampur *et al.*, 2010; Molloy *et al.*, 2010; Xiao and Fayer, 2008).

Further research to determine whether children presenting with diarrhoea may also be co-infected with *Cryptosporidium* spp. and *G. duodenalis* would be useful. But because of wide differences in social class, economic power and culture this may not be important to all of the populations social groupings. The main problem in developing countries is poor sanitation, here *G. duodenalis* infection was common among the age group 4-5 years (Wellington *et al.*, 2008). However, infections can spread rapidly due to the greater social interaction that children have by playing outdoors. Perhaps more importance should be placed on informing local populations of the benefits of good hygiene

and sanitary practices such as hand washing, particularly where people are likely to come into close contact with one another, this maybe the best way of limiting transmission. Increasing levels of sanitation and hygiene are very important to the control of Giardiasis/Cryptosporidiosis infections but a knowledge of a local populations, lifestyle and common practices are also vital to inform risk factors.

This work reinforces the need for handwashing after defecation and the use of boiled water for consumption. Overall, health promotion, public education, improving sanitation conditions, and improving clean drinking water and food are important strategies which can be used for control and prevention of Giardiasis and Cryptosporidiosis infection.

APPENDIX

Appendix 1 Demographic data of symptomatic children from Malawi in relation to *Giardia* and *Cryptosporidium* positive by PCR and DFA

Table 1A. Demographic data of symptomatic children from Karonga in relation to *Giardia* and *Cryptosporidium* positive by PCR and DFA

Study ID #	Sex	DOB	DOC	Age(m)	All positive samples										C-Species
					Giardia					Cryptosporidium sp.					
					RV	DFA	PCR (G-ssu)	G-tpi		PCR (bg)	PCR-gdh	DFA	PCR ₁ (C-ssu)	PCR ₂ (GP60)	
		Date of Birth	Date of collection					PCR	Sequenced						
KAR-010	M	20/11/2005	10/7/2006	8	***	—	+	+	A	+	+	—	—	—	n/a
KAR-029	F	6/1/2005	7/8/2006	17	—	+++++	+	+	B	+	+	—	—	—	n/a
KAR-087	M	27/5/2006	24/10/2006	5	***	—	+	+	B	+	+	—	—	—	n/a
KAR-090	F	21/9/2005	28/10/2006	13	—	—	+	+	nd	+	+	+++	+	+	C. hominis
KAR-094	M	22/2/2006	5/11/2006	9	—	+	+	+	AB	+	+	—	—	—	n/a
KAR-097	F	24/6/2006	8/11/2006	4	***	—	+	+	nd	+	+	—	—	—	n/a
KAR-108	F	5/11/2004	23/11/2006	24	—	—	+	+	nd	—	+	—	—	—	n/a
KAR-110	M	15/5/2005	26/11/2006	6	—	++	+	+	nd	+	+	—	—	—	n/a
KAR-111	F	4/12/2005	4/12/2006	12	—	—	+	+	B	+	+	+++++	+	+	C. parvum
KAR-114	M	25/5/2006	9/12/2006	17	—	++	—	—	nd	—	—	+	—	—	n/a
KAR-118	M	21/4/2006	29/12/2006	8	—	—	—	—	n/a	—	—	+++++	+	+	C. hominis
KAR-120	M	25/2/2006	7/1/2007	10	—	—	+	+	nd	—	+	—	—	—	n/a
KAR-121	M	18/1/2006	9/1/2007	11	—	—	+	+	nd	+	+	—	—	—	n/a
KAR-122	M	31/3/2006	9/1/2007	9	—	—	—	—	n/a	—	—	++	—	—	n/a
KAR-132	F	18/4/2006	21/1/2007	9	—	—	+	+	nd	—	+	+	—	—	n/a
KAR-135	M	21/6/2006	28/1/2007	9	—	—	+	+	nd	+	+	+++++	+	+	C. parvum
KAR-169	F	15/3/2007	1/5/2007	13	*	—	+	+	nd	—	+	—	—	—	n/a
KAR-184	M	11/8/2006	10/6/2007	10	—	—	+	+	nd	+	—	—	—	—	n/a
KAR-190	F	20/11/2005	18/6/2007	19	***	+	—	—	n/a	—	—	—	—	—	n/a
		Positive			4 (26.7%)	5 (14.3%)	15 (23.8%)	15 (23.8%)	5 (33.3%)	11 (17.5%)	14 (22.2%)	7 (20%)	4 (6.3%)	4 (6.3%)	
		Negative			11 (73.3%)	30 (85.7%)	48 (76.2%)	48 (76.2%)	10 (66.7%)	52 (82.5%)	49 (77.8%)	28 (80%)	59 (93.7%)	59 (93.7%)	
Total numbers of samples tested					15	35	63	63	15	63	63	35	63	63	
n = 63 for PCR but n = 35 for DFA															
		Notice;	`* = small numbers of rotavirus (maybe 10 ⁶)					For Crypto		`+ = 1 - 10 oocysts			`++++ = 100 - < 500 oocysts		
			`** = equivalent to positive control							`++ = 10 - < 50 oocysts			`+++++ = > 500 oocysts		
			`*** = large numbers of virus							`+++ = 50 - < 100 oocysts			(uncountable)		

Table 1B. Demographic data of symptomatic children from Lilongwe in relation to *Giardia* and *Cryptosporidium* positive by PCR and DFA

Study ID #	Sex	DOB	DOC	age (m)	All positive samples										C-sp.
					Giardia					Cryptosporidium sp.					
					RV	DFA	PCR (G-ssu)	G-tpi	PCR (B-giardin)	PCR -gdh	DFA	PCR, (C-ssu)	PCR, (GP60)		
		Date of Birth	Date of collection					PCR	Sequenced						
KCH-002	F	27/4/2003	18/7/2005	38	-	-	+	+	B	-	+	+	-	-	n/a
KCH-004	F	13/3/2005	18/7/2005	4	***	-	+	+	AB	-	+	++	+	+	C. hominis
KHC-005	F	24/10/2004	19/7/2005	8	-	-	+	+	AB	-	+	+++	+	-	C. hominis
KCH-014	F	30/6/2005	27/7/2005	2	*	-	-	-	-	-	n/a	+	-	-	n/a
KCH-017	F	30/8/2004	3/8/2005	11	-	-	+	+	AB	+	+	++	+	+	C. hominis
KCH-021	M	19/10/2004	8/8/2005	8	-	-	+	+	AB	-	-	+	+	+	C. parvum
KCH-027	M	22/12/2004	17/8/2005	9	***	-	+	+	A	-	n/a	-	-	-	n/a
KCH-032	M	19/12/2004	22/8/2005	8	***	+	+	+	AB	-	+	-	-	-	n/a
KCH-033	F	15/1/2005	22/8/2005	7	-	+	+	+	AB	-	+	-	+	+	C. hominis
KCH-038	F	1/10/2004	27/8/2005	10	-	-	+	+	AB	-	+	++	+	+	C. hominis
KCH-041	F	7/12/2004	28/8/2005	9	***	-	+	+	AB	-	-	++	-	-	ND
KCH-042	F	12/9/2004	28/8/2005	12	-	-	+	+	B	+	+	++++	+	+	C. hominis
KCH-043	F	24/7/2005	29/8/2005	1	-	-	+	+	A	-	n/a	-	-	-	n/a
KCH-045	F	24/1/2005	29/8/2005	7	**	-	+	+	B	-	n/a	-	-	-	n/a
KCH-051	M	9/1/2005	4/9/2005	7	-	-	+	+	AB	-	n/a	-	-	-	n/a
KCH-053	M	23/1/2005	6/9/2005	8	-	-	+	+	A	-	-	++++	+	+	C. parvum
KCH-054	M	1/3/2005	6/9/2005	6	**	-	+	+	AB	-	+	++	-	-	n/a
KCH-055	F	6/3/2005	6/9/2005	6	***	-	+	+	AB	+	+	+	+	-	C. parvum
KCH-061	F	22/8/2004	10/9/2005	12	-	++	+	+	AB	+	+	++	+	-	C. hominis
KCH-063	M	3/7/2004	10/9/2005	14	-	+	+	+	AB	+	+	+	+	-	C. hominis
KCH-072	F	28/2/2003	12/9/2005	30	-	+	+	+	B	+	+	+	+	+	C. parvum
KCH-078	F	17/10/2004	14/9/2005	11	-	-	+	+	B	+	+	++	+	+	C. hominis
KCH-081	M	17/5/2002	15/9/2005	37	-	-	+	+	AB	-	+	+	+	+	C. parvum
RV Co-infection with Giardia					RV Co-infection with Crypto										
Positive					7 (31.8%)	5 (6.9%)	22 (30.6%)	22 (30.6%)	22 (30.6%)	7 (9.7%)	15 (20.8%)	17 (23.6%)	14 (19.4%)	10 (13.9%)	5 (27.8%)
Negative					15 (68.2%)	67(93.1%)	50 (69.5%)	50 (69.5%)	50 (69.5%)	65 (90.3%)	57 (79.2%)	55 (76.4%)	58 (80.6%)	62 (86.1%)	13 (72.2%)
Total numbers of samples tested					22	72	72	72	72	72	72	72	72	72	18
n = 72 for both PCR & DFA					For Cryptosporidium										
** = small numbers of rotavirus (maybe 10 ⁶)										'++ = 10 - < 50 oocysts					
*** = equivalent to positive control										'++++ = > 500 oocysts					
**** = large numbers of virus										'+++ = 50 - < 100 oocysts (uncountable)					

Table 1C. Demographic data of symptomatic children from Mangochi in relation to *Giardia* and *Cryptosporidium* positive by PCR and DFA

Study ID #	Sex	DOB	DOC	age (m)	All positive samples										RV Co-infection with <i>Crypto</i>
					<i>Giardia</i>					<i>Cryptosporidium</i> sp.					
					RV	DFA	PCR (G-ssu)	G-tpi	Sequenced	PCR (B-giardin)	PCR - gdh	DFA	PCR, (C-ssu)	PCR, (GP60)	
		Date of Birth	Date of collection					PCR							
MAN-010	F	4/12/2005	14/7/2006	11	**	-	+	+	A	-	n/a	+	-	n/a	n/a
MAN-013	M	25/12/2005	24/7/2006	7	-	-	+	+	nd	+	n/a	-	-	n/a	n/a
MAN-015	M	14/12/2005	27/7/2006	7	-	-	+	+	A	-	+	+	-	n/a	n/a
MAN-021	M	8/10/2005	13/8/2006	10	-	-	+	+	A	-	-	-	-	n/a	n/a
MAN-029	M	1/1/2006	30/8/2006	7	-	-	+	+	nd	-	n/a	+++	+	+	<i>C. hominis</i>
MAN-030	M	22/2/2006	30/8/2006	7	-	-	+	+	A	-	+	+	-	n/a	n/a
MAN-035	F	5/6/2003	6/9/2006	38	-	-	+	+	nd	-	+	-	-	n/a	n/a
MAN-051	M	20/1/2006	15/11/2006	10	-	-	+	+	nd	-	n/a	++++	+	+	<i>C. hominis</i>
MAN-053	M	2/12/2004	23/11/2006	23	-	-	+	+	AB	-	+	-	-	n/a	n/a
MAN-064	F	5/6/2006	18/12/2006	6	-	-	+	+	nd	-	+	-	-	n/a	n/a
MAN-067	M	26/10/2006	19/12/2006	8	-	+	+	+	nd	+	+	-	-	n/a	n/a
MAN-076	F	30/4/2006	3/1/2007	8	-	-	+	+	A	-	+	+	-	n/a	n/a
MAN-079	M	30/12/2005	17/1/2007	12	***	+	+	+	nd	-	+	-	-	n/a	n/a
MAN-081	F	1/7/2006	15/12/2006	5	-	-	+	+	nd	-	n/a	-	-	n/a	n/a
MAN-087	F	17/2/2006	4/2/2007	12	-	-	+	+	nd	-	-	-	-	n/a	n/a
MAN-107	F	4/10/2006	27/3/2007	5	-	-	+	+	nd	-	+	+	-	n/a	n/a
MAN-115	F	4/3/2006	5/4/2007	11	**	+	-	-	n/a	-	n/a	-	-	n/a	n/a
MAN-135	M	21/4/2005	17/5/2007	25	**	+	+	+	AB	-	+	-	-	n/a	n/a
MAN-137	M	13/3/2005	23/5/2007	26	-	+	+	+	B	-	-	+	-	n/a	n/a
MAN-144	M	20/11/2006	13/6/2007	19	-	-	+	+	nd	-	+	-	-	n/a	n/a
				RV Co-infection with <i>Giardia</i>											
		Positive			2 (10.5%)	5 (20%)	19 (29.7%)	19 (29.7%)	8 (40%)	2 (3.1%)	11 (17.2%)	8 (32%)	2 (3.1%)	2 (3.1%)	1 (12.5%)
		Negative			17 (89.5%)	20(80%)	45 (70.3%)	45 (70.3%)	12 (60%)	62 (96.9%)	53 (82.8%)	17 (64%)	62 (96.9%)	62 (96.9%)	7 (87.5%)
Total numbers of samples tested					19	25	64	64	20	64	64	25	64	64	8
n = 64 for PCR but n = 25 for DFA															
		Notice;	* = small numbers of rotavirus (maybe 10 ⁵)					For <i>Cryptosporidium</i>					* = 1 - 10 oocysts		* = 100 - < 500 oocysts
			** = equivalent to positive control										* = 10 - < 50 oocysts		* = > 500 oocysts
			*** = large numbers of virus										* = 50 - < 100 oocysts		(uncountable)

Appendix 2. The distribution of assemblage A and B based on the different loci and mixed infection

Table 2A. The distribution of *Giardia* assemblage A and B from Malawi based on the different loci and mixed infection

Subject no.	Sample ID	PCR Positivity			<i>Tpi</i> gene		<i>Bg</i> gene		<i>Gdh</i> gene	
		Ass. A	Ass. B	Assemblage	Sub-assemblage	Accession No.	Sub-assemblage	Accession No.	Sub-assemblage	Accession No.
1	K10	-	+	B	-	-	AI	KM190690	-	-
2	K29	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV	KC96064.1
3	K87	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)
4	K90	-	+	B	BIII	AF069561/ BAH-12 (Australia)	BIII	JF918485	-	-
5	K94	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	Other B	KC632652 (Thailand)	BIV (H43)	EF507682 (Brazil)
6	K97	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	-	-
7	K108	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	-	-
8	K110	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	B Heterogenous	JF918489 (India)	-	-
9	K111	+	+	AB	BIV (VB960855) Other B	KM190840 (Canada) KF843920 (Germany)	AI	KM190690	BIV	KC96064.1

10	K121	-	+	B	All & BIV	KR260616 (Egypt) AF069560 (Australia)	B Heterogenous	JF918489 (India)	BIV (H43)	EF507682 (Brazil)
11	K122	+	-	A	All	KR260616 (Egypt)	-	-	-	-
12	K135	-	+	B	-	-	-	JF918489 (India)	BIV (H43)	EF507682 (Brazil)
13	K184	-	+	B	-	-	B Heterogenous	JF918489 (India)	-	-
14	L2	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)
15	L4	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)
16	L5	+	+	AB	All	KR260616 (Egypt)	-	-	BIV (KC96064.1) BIV (H43)	KC96064.1 EF507682 (Brazil)
17	L21	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	-	-	-	-
18	L27	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	-	-
19	L32	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	-	-	BIV BIV (H43)	KC96064.1 EF507682 (Brazil)
20	L33	+	+	AB	All & BIV	KR260616 (Egypt) AF069560 (Australia)	-	-	BIV BIV (H43)	KC96064.1 EF507682 (Brazil)

21	L38	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/BAH-12 (Australia)	-	-	-	-
22	L41	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	-	-	-	-
23	L42	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)
24	L54	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	-	-	BIV BIV (H43)	KC96064.1 EF507682 (Brazil)
25	L55	-	+	B	-	-	B (Heterogenous)	JF918489 (India)	BIV BIV (H43)	KC96064.1 EF507682 (Brazil)
26	L61	-	+	B	-	-	-	-	BIV	KC96064.1
27	L63	-	+	B	BIII	AF069561/ BAH-12 (Australia)	Other B	KC632652 (Thailand)	BIV BIV (H43)	KC96064.1 EF507682 (Brazil)
28	L72	-	+	B	-	-	BIII	JF918485 (India)	BIV BIV (H43) BIV Cla145)	KC96064.1 EF507682 (Brazil) HM134212.1
29	L78	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	All	JX898210 (China)	-	-
	L81	+		AB	All		-		BIV (H43)	EF507682 (Brazil)

30			+			KR260616 (Egypt)		-	BIV (Cla145)	HM134212.1 (Brazil)
31	M13	-	+	B	-	-	-	-	BIV (H43)	EF507682 (Brazil)
31	M30	-	+	B	-	-	-	-	BIV)	KC96064.1
32	M35	-	+	B	-	-	-	-	BIV (H43)	EF507682 (Brazil)
33	M53	-	+	B	-	-	-	-	BIV (Cla145)	HM134212.1 (Brazil)
34	M64	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)
35	M67	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	Other B	KC632652 (Thailand)	BIV (H43)	EF507682 (Brazil)
36	M76	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV	KC96064.1
37	M79	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)
38	M87	+	-	A	All	KR260616 (Egypt)	-	-	-	-
39	M107	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)
40	M135	+	+	AB	All & BIII	KR260616 (Egypt) AF069561/BAH-12 (Australia)	-	-	-	-

41	M137	+	+	AB	AII & BIII	KR260616 (Egypt) AF069561/BAH-12 (Australia)	-	-	-	-
42	M144	-	+	B	BIII	AF069561/ BAH-12 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)

Table 2B. The distribution of *Giardia* assemblage A and B from Cambodia showing all positive results by PCR and based on the different loci and mixed infection

Subject no.	Sample ID	PCR Positivity			<i>Tpi</i> gene		<i>Bg</i> gene		<i>Gdh</i> gene	
		Ass. A	Ass. B	<i>Assemblage</i>	Sub-assemblage	Accession No.	Sub-assemblage	Accession No.	Sub-assemblage	Accession No.
1	B10	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 (Australia) KP687792 (Canada)	BG10 = BIII	JF918485 JF918500	-	-
2	A15 & B15	+	+	AB	BIV (VB960855)	KM190840 (Canada) KC632562 (Australia) KP687792 (Canada)	-	JF918480 (India)	BIV (H43)	EF507682 (Brazil)
3	A19	+	-	A	AII	KR260616 (Egypt)	-	-	-	-
4	B27	-	+	B	-	-	BG27 = BIII	JF918500	-	-
5	A28	+	-	A	AII	KR260616 (Egypt)	-	-	-	-
6	B42	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 (Australia) KP687792 (Canada)	BG42 = B Heterogenous	JF918489 (India)	GD42 = B (VAN/94/UB C/122)	VAN/94/UBC/122
7	B56	-	+	B	BIII	KC632557 (Australia)	-	-	BIV (H43)	EF507682 (Brazil)

8	B67	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	-	-	GD67 = B (VAN/90/UB C.54)	KP687770 (Canada; Creston)
9	B68	-	+	B	BIII	KC632557 (Australia)	-	-	-	-
10	B75	-	+	B	BIII	KC632557 (Australia)	-	-	-	-
11	A79 & B79	+	+	AB	All & BIII	KR260616 (Egypt) KC632557 (Australia)	BG79 = All	KM0698	All	AY178737 (Bris-136) (Australia)
12	B82	-	+	B	Other B	KF843920 (Germany)	BG82 =BIII	JF918500 (India)	GD82 = BIV (H43)	EF507682 (Brazil)
13	B84	-	+	B	B heterogenous & BIII	GU564279 (China) KC632557 (Australia)	BG84 =B Heterogenous & Other B	JF918489 (India) JX994239.1 (China)	GD84 = B (VAN/90/UB C.54)	KP687770 (Canada; Creston)
14	B102	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	-	-	-	-
15	B110	-	+	B	BIII	KC632557 (Australia)	-	-	-	-
16	B121	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia)	-	-	GD121 = B (VAN/90/UB C.54)	

						KP687792 (Canada)				KP687770 (Canada; Creston)
17	A155	+	-	A	All	KR260616 (Egypt)	BG155 = AI	KX302199	AI	M84604 (Portland 1)
18	B174	-	+	B	-	-	-	-	GD174 = B (VAN/94/UB C/122)	VAN/94/UBC/1 22
19	B257	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	BG257 = Other B	JX994239 China	GD257 = BIV	KC96064.1
20	A272	+	-	A	All	U57897 (United State of America)	-	-	All	AY178737 (Bris-136) (Australia)
21	B275	-	+	B	-	-	-	-	GD275 = BIV	KC96064.1
22	B293	-	+	B	-	-	-	-	GD293 = BIV (H43)	EF507682 (Brazil)
23	B339	-	+	B	-	-	-	-	GD339 = BIV	KC96064.1
24	B371	-	+	B	-	-	-	-	GD293 = BIV (H43)	EF507682 (Brazil)
25	A411 & B411	+	+	AB	BIII	KC632557 (Australia) -	BG411 = AI	KY499042 (Spain)	-	-
26	B439	-	+	B	-	-	-	-	GD439 = BIV (H43)	EF507682 (Brazil)
27	B460	-	+	B	-	-	-	-	-	-

28	B467	-	+	B	-	-	-	-	GD467 = BIV	KC96064.1
29	B494	-	+	B	-	-	-	-	-	-
30	A530	+	-	A	All	KR260616 (Egypt)	-	-	All	AY178737 (Bris-136) (Australia)
31	B549	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	-	-	GD549 = BIV	KC96064.1
32	B558	-	+	B	-	-	-	-	BIV (H43)	EF507682 (Brazil)
33	B591	-	+	B	-	-	-	-	GD591 = B (VAN/90/UB C.54)	KP687770 (Canada; Creston)
34	B612	-	+	B	-	-	-	-	BIV (H43)	EF507682 (Brazil)
35	B646	-	+	B	-	-	-	-	-	-
36	B656	-	+	B	-	-	-	-	GD656 = BIV	KC96064.1
37	B658	-	+	B	-	-	-	-	GD658 = BIV	KC96064.1
38	B669	-	+	B	-	-	-	-	GD669 = BIV	KC96064.1
39	B697	-	+	B	-	-	-	-	-	-
40	B702	-	-	B	BIII	KC632557 (Australia)	-	-	GD702 = BIV (H43)	EF507682 (Brazil)
41	B709	-	+	B	-	-	-	-	-	-

42	B714	—	+	B	-	-	-	-	GD714 = BIV (H43)	EF507682 (Brazil)
43	B727	—	+	B	-	-	-	-	GD727 = B (VAN/94/UB C/122)	VAN/94/UBC/1 22
44	B729	—	+	B	-	-	-	-	-	-
45	B734	—	+	B	-	-	-	-	-	-
46	B735	—	+	B	-	-	-	-	-	-
47	B738	—	+	B	-	-	-	-	GD738 = B (VAN/90/UB C/54)	KP687770 (Canada; Creston)
48	B739	—	+	B	-	-	-	-	-	-
49	B747	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	-	-	GD747 = BIV	KC96064.1
50	B749	-	+	B	BIII	KC632557 (Australia)	-	-	GD749 = BIV	KC96064.1
51	B761	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	BG761 = B (Heterogenous)	JF918489 (India)	GD761 = B (VAN/94/UB C/122)	VAN/94/UBC/1 22
52	B780	—	+	B	B (heterogenous) & BIII	GU564279 (China) KC632557 (Australia)	-	-	-	-
53	B802	-	+	B	BIII	KC632557 (Australia)	-	-	-	-

54	B817	-	+	B	-	-	BG817 = BIII	JF918500	-	-
55	B818	-	+	B	-	-	-	-	-	-
56	A824	+	-	A	-	-	-	-	All	AY178737 (Bris-136) (Australia)
57	A843	+	-	A	-	-	BG843 = All KM190698	JF918500	-	-
58	B882	-	+	B	-	-	-	-	GD882 = BIV	KC96064.1
59	B884	-	+	B	-	-	-	-	All	AY178737 (Bris-136) (Australia)
60	A888	+	-	A	All	KR260616 (Egypt)	-	-	All	AY178737 (Bris-136) (Australia)
61	B907	-	+	B	-	-	-	-	GD907 = BIV (H43)	EF507682 (Brazil)
62	B977	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	BG977 = BIII	JF918500	GD977 = B (VAN.94/UB C/122)	VAN.94/UBC/1 22
63	B987	-	+	B	-	-	-	-	GD987 = BIV (H43)	EF507682 (Brazil)
64	B989	-	+	B	-	-	BG989 = BIII	JF918500	GD989= BIV	KC96064.1
65	B1001	-	+	B	B (heterogenous) & BIII	GU564279 (China) KC632557 (Australia)	BG1001 = B (Heterogenous)	JF918489 (India)	GD1001= BIV	KC96064.1

66	B1011	-	+	B	BIV (VB960855)	KM190840 (Canada) KC632562 Australia) KP687792 (Canada)	BG1011 = Other B	JX994239 China	GD1011 = BIV (H43)	EF507682 (Brazil)
67	B1032	-	+	B	-	-	BG761 = B (Heterogenous)	JF918489 (India)	-	-
68	B1168	-	+	B	-	-	-	-	GD1168 = B (VAN/90/UB C/54)	KP687770 (Canada; Creston)
69	A1074 & B1074	+	+	AB	All & BIII	KR260616 (Egypt) KC632557 (Australia)	-	-	-	-
70	B1092	-	+	B	-	-	-	-	-	-
71	B1110	-	+	B	-	-	-	-	-	-
72	B1134	-	+	B	-	-	-	-	GD1134 = B (VAN/94/UB C/122)	VAN/94/UBC/1 22
73	B1139	-	+	B	-	-	-	-	-	-
74	B1168	-	+	B	-	-	-	-	-	-

Table 2C. Demographic data of symptomatic children from Cambodia in relation to the distribution of *Giardia* sub-assemblage B based on the different loci and mixed infection

Subject no.	gender	Province	location	<i>tpi</i>	<i>bg</i>	<i>gdh</i>
10	F	Siem Reap	AHC OPD	BIV VB960855	BIII	-
15	M	Siem Reap	AHC OPD	BIV VB960855	Other B	BIV H43 (EF507682)
42	M	Siem Reap	AHC OPD	BIV VB960855	B Heterogenous	B VAN/94/UBC/122
67	F	Banteay Meanchey	AHC IPD	BIV VB960855	-	B VAN/90/UBC/54
82	M	Siem Reap	AHC OPD	Other B	BIII	BIV H43 (EF507682)
84	M	Battambang	AHC OPD	B Heterogenous	Other B	B VAN/90/UBC/54
102	F	Oddar Meanchey	AHC OPD	BIV VB960855	-	-
121	F	Siem Reap	AHC IPD	BIV VB960855	-	B VAN/90/UBC/54
257	F	Siem Reap	AHC OPD	BIV VB960855	Other B	BIV KC96064.1
549	M	Siem Reap	AHC IPD	BIV VB960855	-	BIV KC96064.1
747	M	Siem Reap	AHC OPD	BIV VB960855	-	BIV KC96064.1
977	M	Siem Reap	AHC OPD	BIV VB960855	BIII	-
1001	F	Siem Reap	AHC OPD	B Heterogenous	B Heterogenous	BIV KC96064.1
1011	M	Siem Reap	AHC OPD	BIV VB960855	Other B	BIV H43 (EF507682)

Appendix 3. Sequence chromatograms of nucleotide variations. Chromatogram of sequences generated from *Giardia* Assemblage A and B using multi-locus markers.

Figure 3A. Chromatogram of a sequence generated from *Giardia* Assemblage B using *tpi* gene (forward primer)



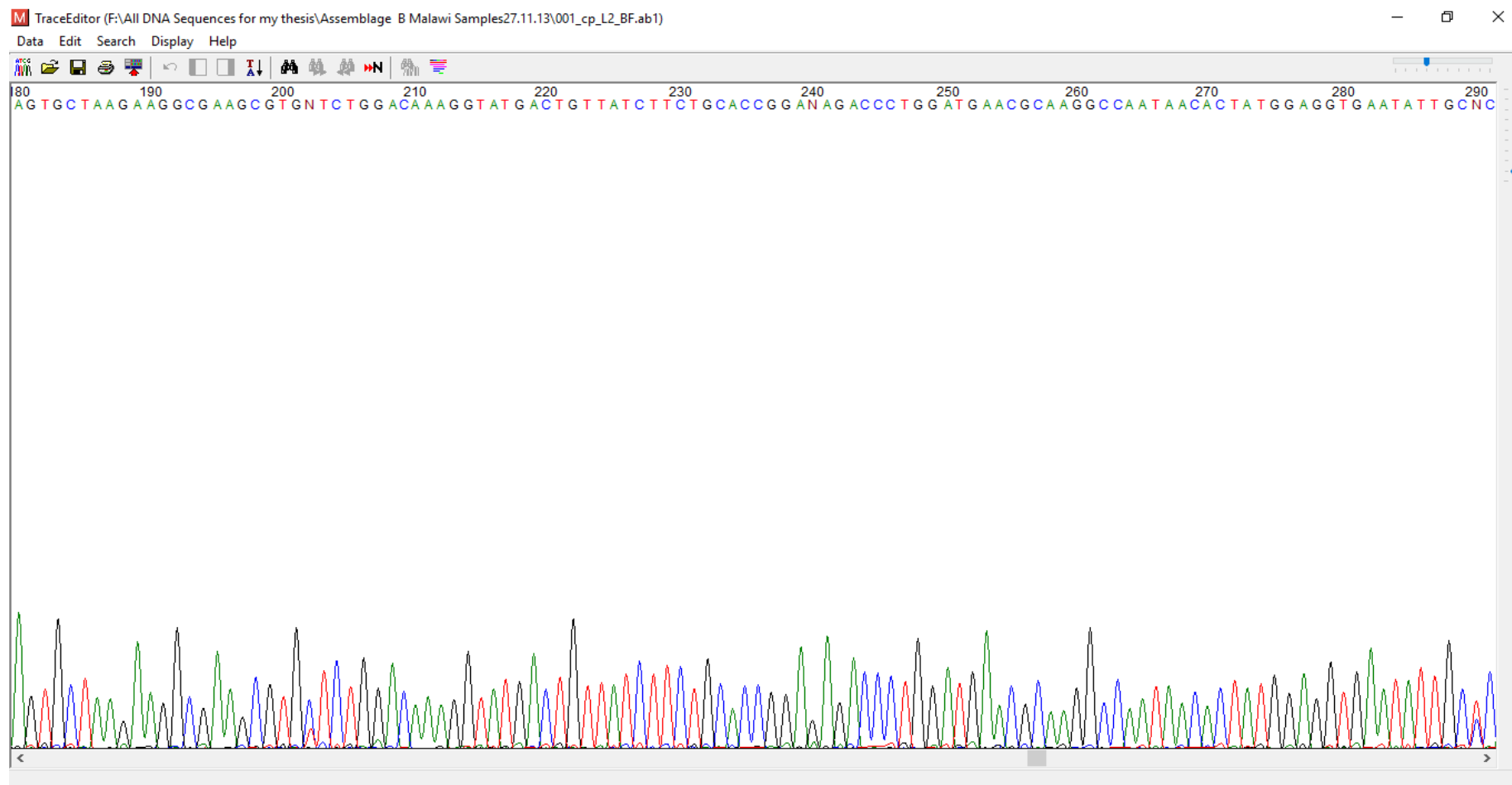
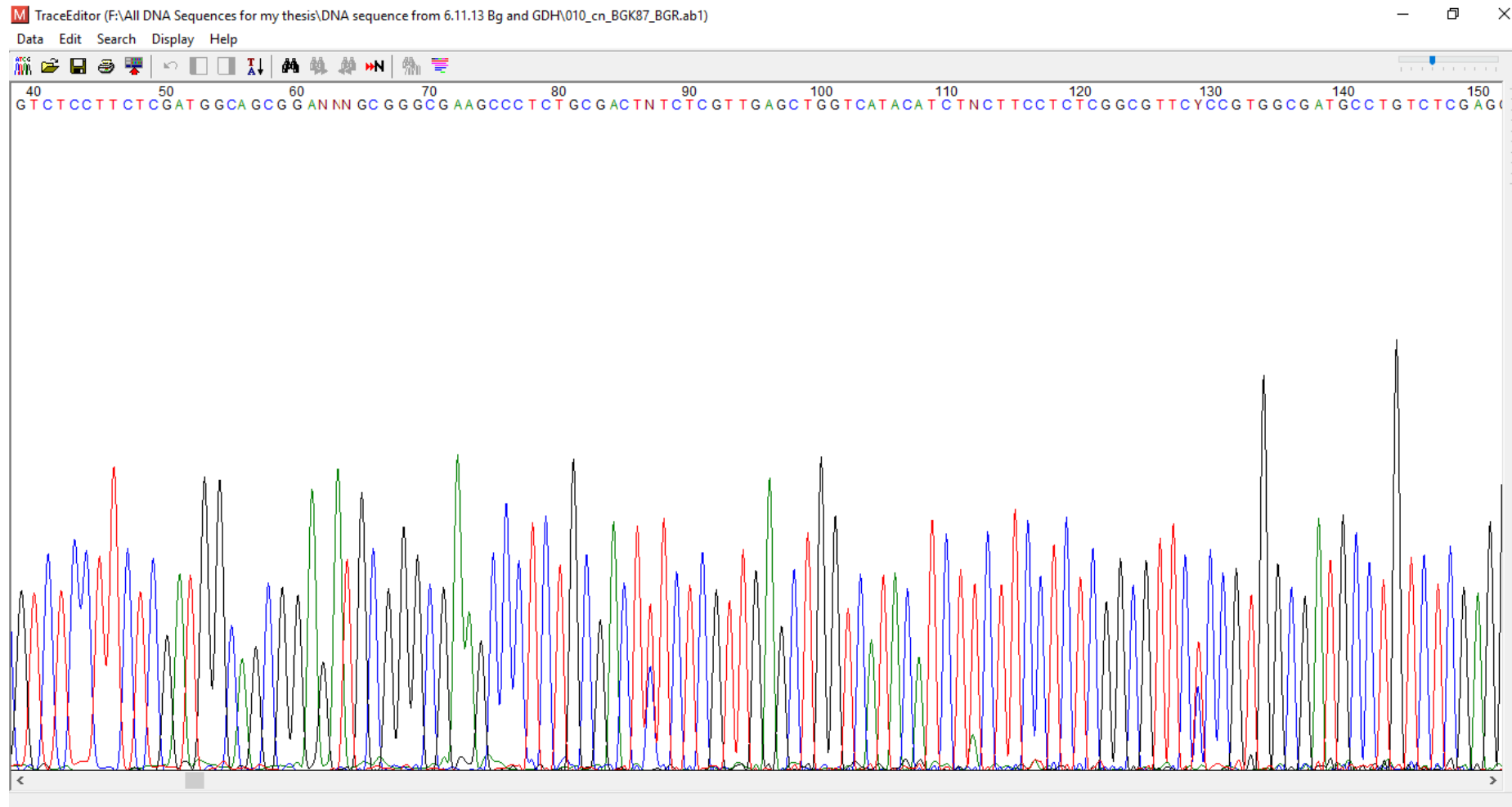
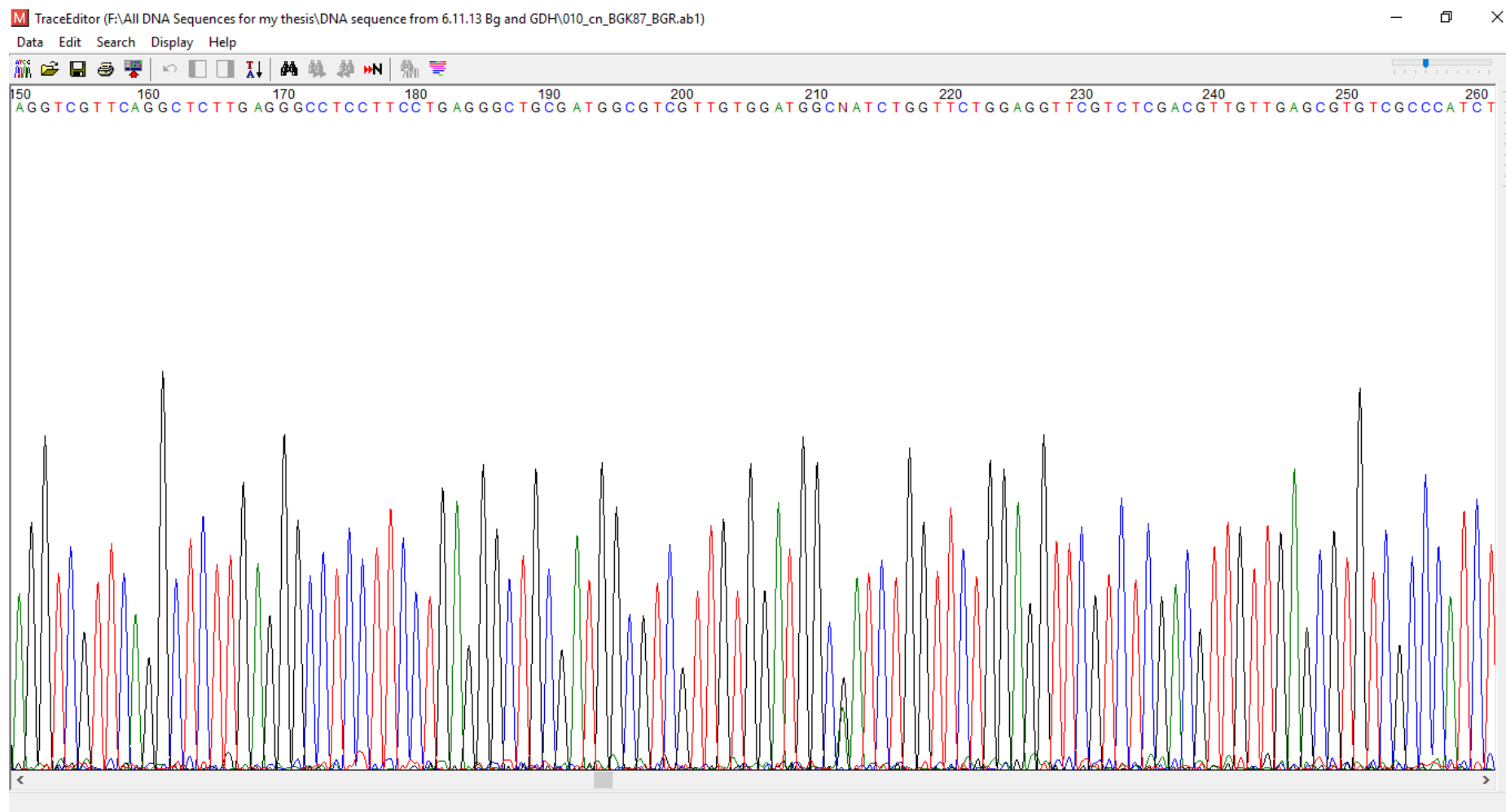
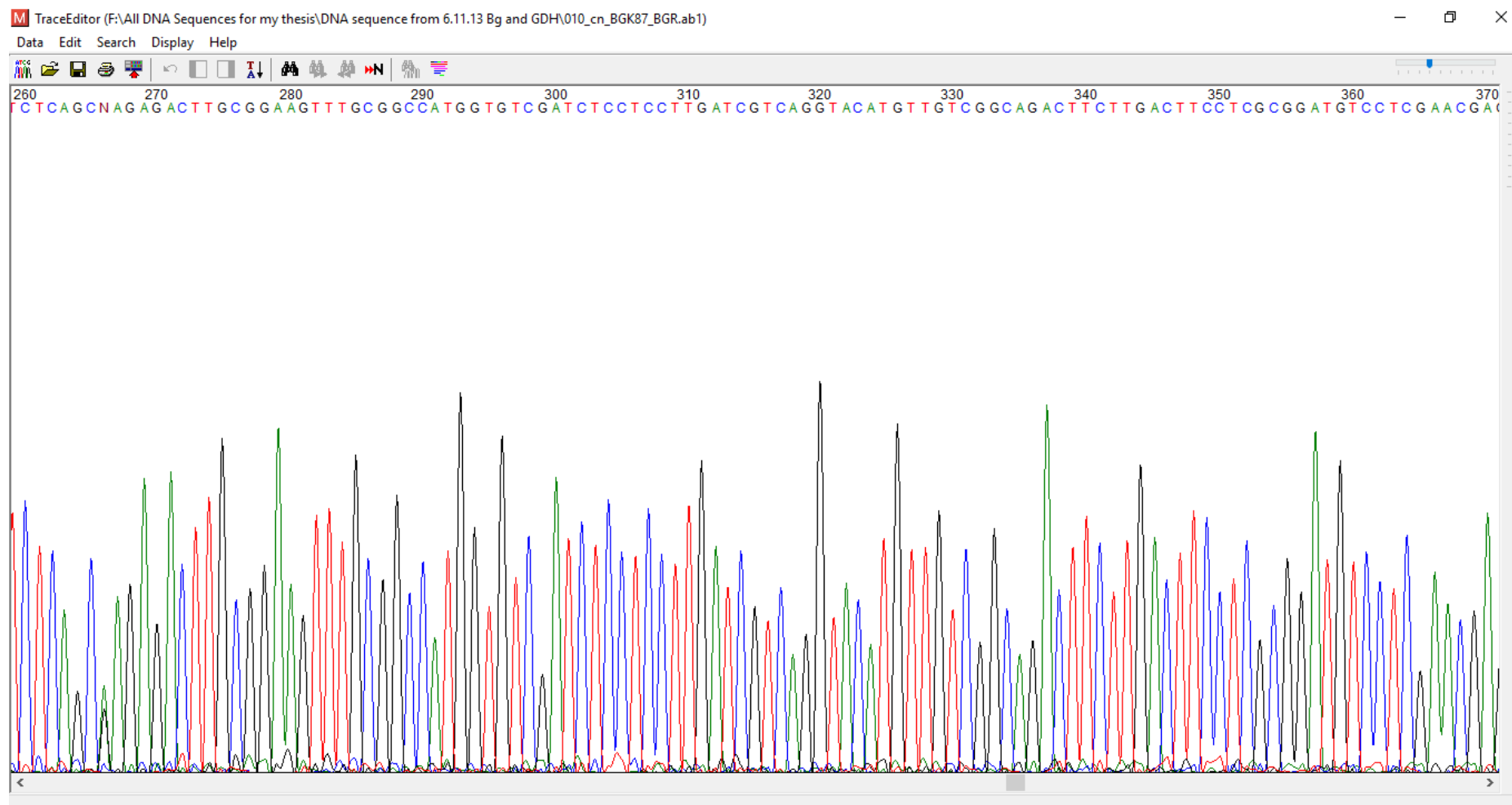




Figure 3B. Chromatogram of a sequence generated from *Giardia* Assemblage A and B using *bg* gene (reward primer)







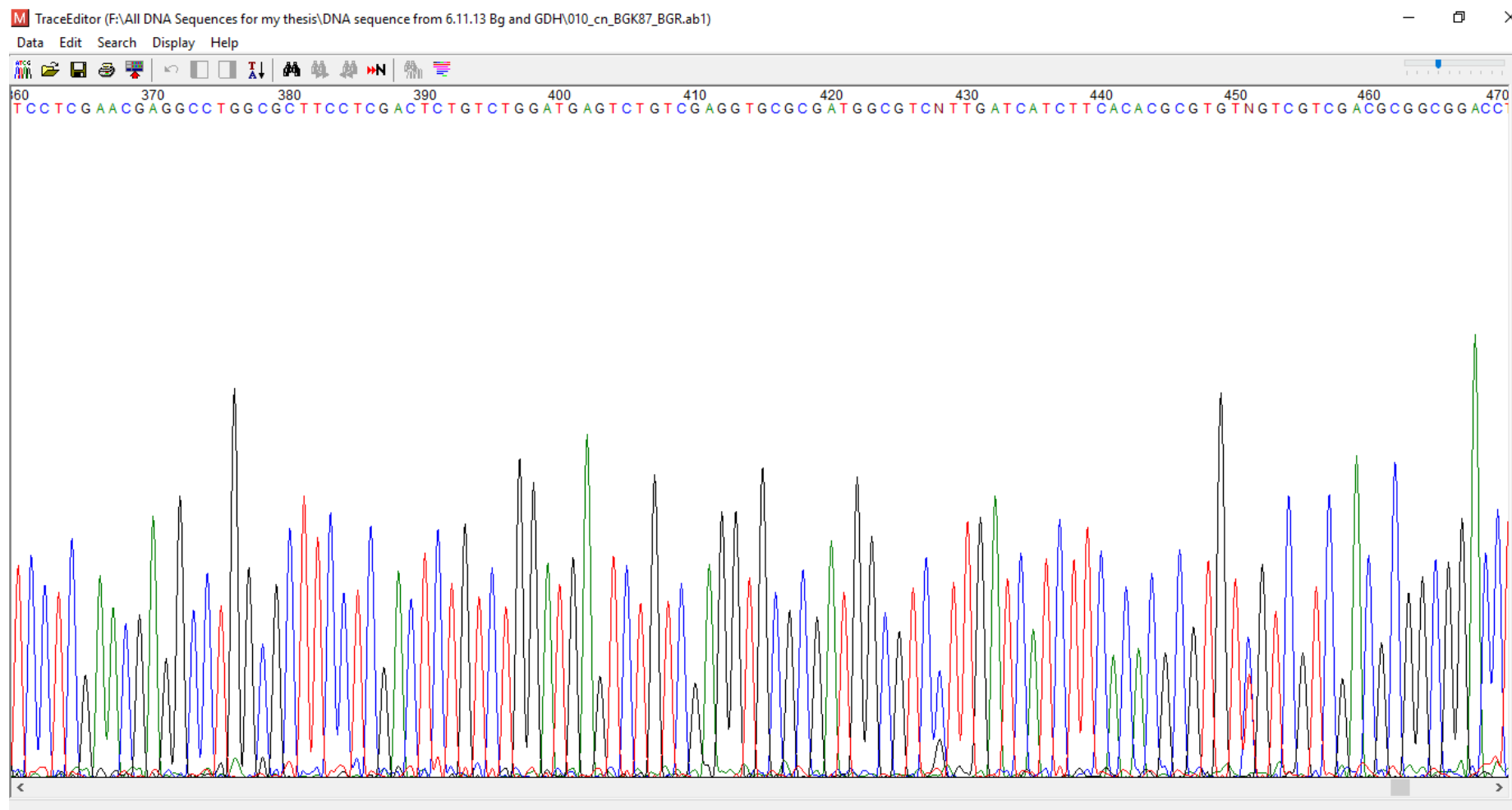
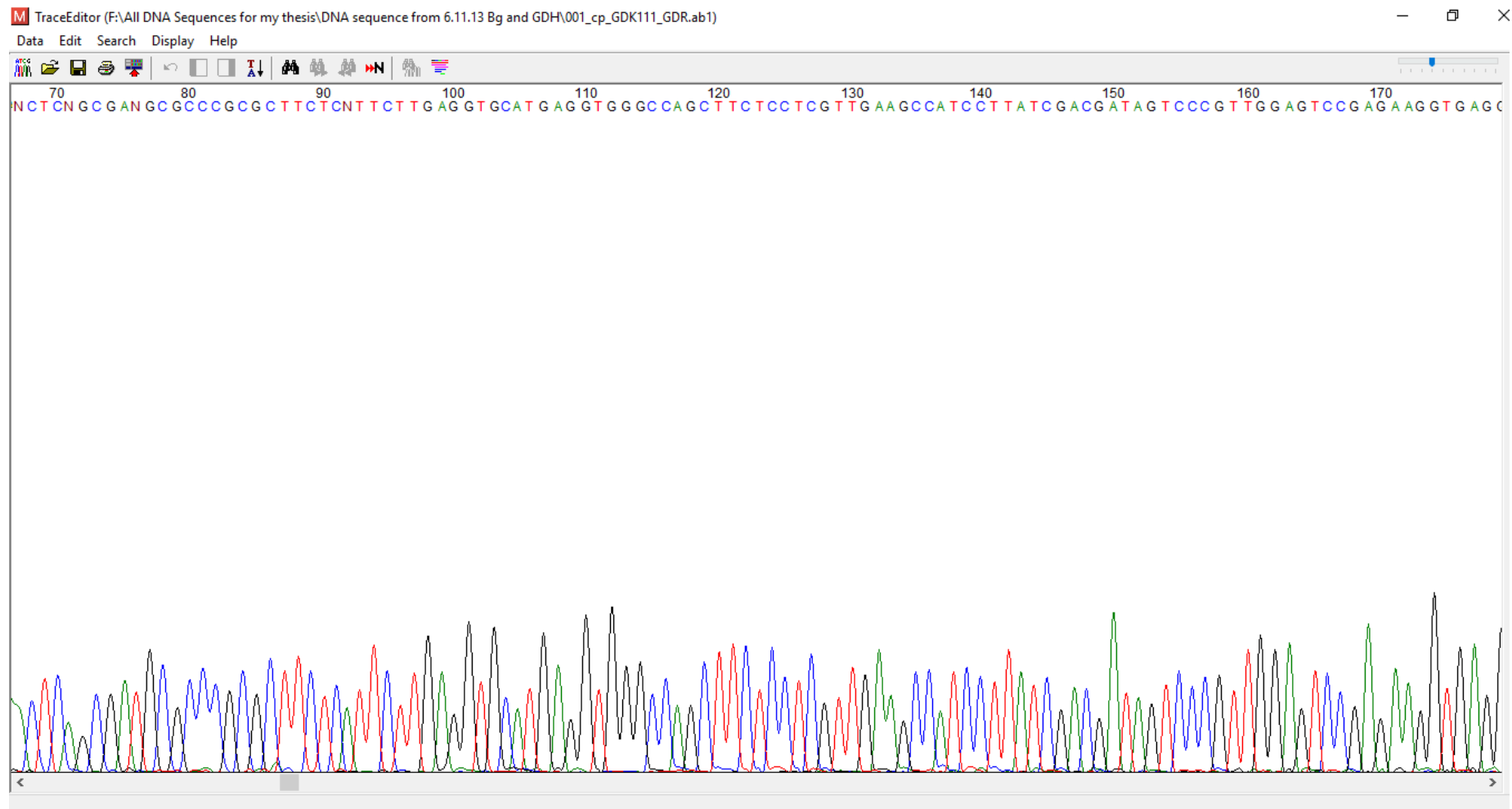
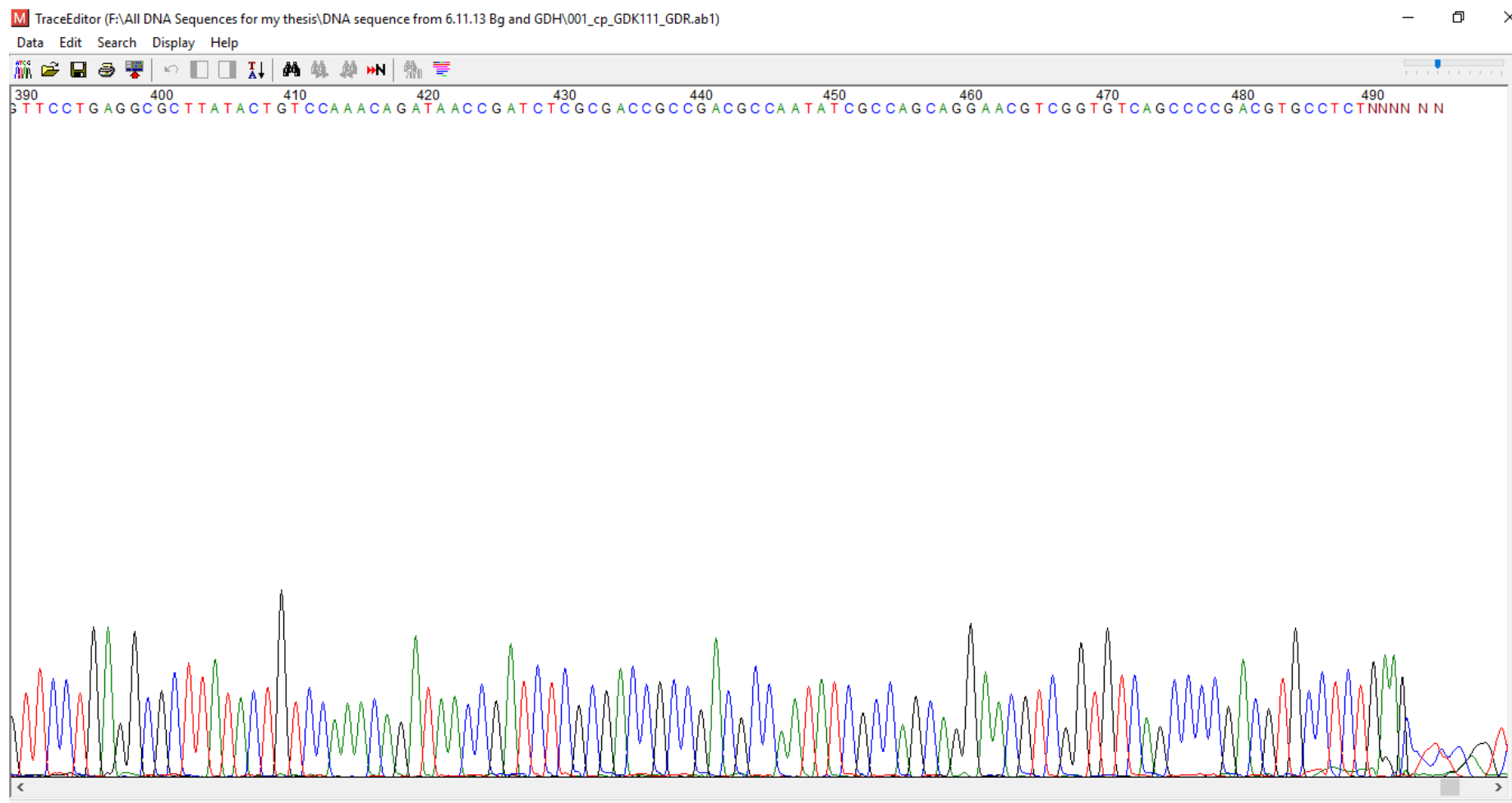


Figure 3C. Chromatogram of a sequence generated from *Giardia* Assemblage A and B using *gdh* gene (reward primer)









Appendix 4. Multiple alignments of gene (*tpi*, *bg* and *gdh*) sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B.

Figure 4A. Multiple alignments of *tpi* sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B.

CLUSTAL O(1.2.4) multiple sequence alignment

```

A530      ACGTCCTTGCCAAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
A155      AGCTCCTTGCCAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
A79       AGCTCCTTGCCAAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
AK121     AGCTCCTTGCCAAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
AL78      AGCTCCTTGCCAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
AM137     AGCTCCTTGCCAAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
A28       AGCTCCTTGCCAAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
U57897_Ref._AII      AGCTCCTTGCCAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
DQ650648_Ref._AIII   AGCTCCTTGCCAAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
L02120_Ref._AI_isolate_WB  AGCTCCTTGCCAAGCGCCTCAAGCTGGGCGATGTTACCTCCATGGTGCGGTTGGCCTTG      60
*          ****

A530      CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
A155      CGCTCACCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
A79       CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
AK121     CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
AL78      CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
AM137     CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
A28       CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
U57897_Ref._AII      CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
DQ650648_Ref._AIII   CGCTCATCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
L02120_Ref._AI_isolate_WB  CGCTCGTCCAAAGGTCTCTCCGACGCGAGAAGATGACCGTCATCCCCTTTTCCAGGGCACGC      120
*****

A530      TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
A155      TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
A79       TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
AK121     TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
AL78      TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
AM137     TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
A28       TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
U57897_Ref._AII      TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
DQ650648_Ref._AIII   TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
L02120_Ref._AI_isolate_WB  TTAGCCTTCTTGCGCTTTGCTCGTCGGTCTCCCCCATGATTCTGCGTCTTTTCAGAGTGC      180
*****

```

A530	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCGCCAGTC	240
A155	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACGGGTCTCGCCAGTC	240
A79	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACGTGTCTCGCCAGTC	240
AK121	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCGCCAGTC	240
AL78	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCGCCAGTC	240
AM137	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCGCCAGTC	240
A28	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCGCCAGTC	240
U57897_Ref._AII	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCGCCAGTC	240
DQ650648_Ref._AIII	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCACCAGTC	240
L02120_Ref._AI_isolate_WB	CCTACTATCACATGCTTCAAACCCATGTCCTGAAGCATCTCAACACTTGTCTCGCCAGTC	240

*****.*****

A530	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
A155	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
A79	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
AK121	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
AL78	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
AM137	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
A28	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
U57897_Ref._AII	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287
DQ650648_Ref._AIII	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCCGCTATCCTCAA	287
L02120_Ref._AI_isolate_WB	CACGCCCCGTTCCCCTCTAGGTACACATTCTGCGCTGCTATCCTCAA	287

CLUSTAL O(1.2.4) multiple sequence alignment

```

BL33          ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTTACCTCCATAGTGTTATTGGCCTTG 60
AF069560_Ref._BIV_Ad-19 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTTACCTCCATAGTGTTATTGGCCTTG 60
B82           ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
B549          ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
B780          ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
B10           ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
BK111         ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
KF843920_isolate_VB906855 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
B15           ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
BM64          ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
B56           ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
BL4           ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
BK29          ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
AF069561_Ref._BIII_isolate_BAH-12 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
KC632557_Ass_B_isolate_HS29 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
KC632558_isolate_HS114 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
BM79          ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
HM140722_isolate_Sweh171 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
HM140714_Ass_B_Sweh060 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
BM137         ATCTCCTTCTTAAGAGCCTCGAGCTGGGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
GU564279_isolate_RM1 ATCTCCTTCTTAAGAGCCTCGAGCTGAGCAATATTCACCTCCATAGTGTTATTGGCCTTG 60
*****_*****

```


BL33	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
AF069560_Ref._BIV_Ad-19	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
B82	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
B549	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
B780	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
B10	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
BK111	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
KF843920_isolate_VB906855	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
B15	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
BM64	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
B56	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
BL4	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
BK29	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
AF069561_Ref._BIII_isolate_BAH-12	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
KC632557_Ass_B_isolate_HS29	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
KC632558_isolate_HS114	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
BM79	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
HM140722_isolate_Sweh171	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
HM140714_Ass_B_Sweh060	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
BM137	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120
GU564279_isolate_RM1	CGTTCATCCAGGGTCTCTCCGGTGCAGAAAGATAACAGTCATACCTTTGTCCAGAGCACGC	120

** *****_*****

BL33	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
AF069560_Ref._BIV_Ad-19	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
B82	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
B549	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
B780	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
B10	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
BK111	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
KF843920_isolate_VB906855	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
B15	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
BM64	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
B56	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
BL4	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
BK29	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
AF069561_Ref._BIII_isolate_BAH-12	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
KC632557_Ass_B_isolate_HS29	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
KC632558_isolate_HS114	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
BM79	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
HM140722_isolate_Sweh171	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
HM140714_Ass_B_Sweh060	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
BM137	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180
GU564279_isolate_RM1	TTCGCCCTTCTTAGCACTCTGCTCATTGGTCTCGCCCATGATTCTACGTCTTTCAGAGTGT	180

*****_*****

BL33	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
AF069560_Ref._BIV_Ad-19	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
B82	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
B549	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
B780	CCTACTATTACATAGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
B10	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
BK111	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
KF843920_isolate_VB906855	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
B15	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
BM64	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
B56	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
BL4	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
BK29	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
AF069561_Ref._BIII_isolate_BAH-12	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
KC632557_Ass_B_isolate_HS29	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
KC632558_isolate_HS114	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
BM79	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
HM140722_isolate_Sweh171	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
HM140714_Ass_B_Sweh060	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
BM137	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240
GU564279_isolate_RM1	CCTATTATTACATGGCTCAGCCCCATGTCCAGCAGCATCTCGACGCTTGTCTCGCCGGTC	240

*** :***** _***** *****

BL33	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
AF069560_Ref._BIV_Ad-19	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
B82	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
B549	CATGCACCATTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
B780	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
B10	CATGCACCATTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
BK111	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
KF843920_isolate_VB906855	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
B15	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
BM64	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
B56	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
BL4	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
BK29	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
AF069561_Ref._BIII_isolate_BAH-12	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
KC632557_Ass_B_isolate_HS29	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
KC632558_isolate_HS114	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
BM79	CAAGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
HM140722_isolate_Sweh171	CATGCACCGTTCCCTTCCAGATACACATTCTGTGCTGCTATTTTTCAG	287
HM140714_Ass_B_Sweh060	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287
BM137	CATGCACCGTTCCCTTCCAGATACACATTCTGTGCTGCTATTTTTCAG	287
GU564279_isolate_RM1	CATGCACCGTTCCCTTCCAGATACACGTTCTGTGCTGCTATTTTTCAG	287

*** :***** _***** *****

Figure 4B. Multiple alignments of *bg* sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B.

CLUSTAL O(1.2.4) multiple sequence alignment

BG79	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BG1032	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BG254	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
JX898210	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
HM165226	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
AY072728	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
DQ090527	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
DQ090522	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
AY072727	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
HM165213	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
HM165208	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
HM165216	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BG257	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BG42	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BG15	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
HM165214	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
AY072725	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
JQ303245	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
X85958	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
AY072723	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
DQ090523	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BGM13	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BGL78	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BGK10	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BGK111	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BG411	TAAACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BGL63	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BGK110	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGCGGCGGAGATGCGGGCGAAGC	60
BGM67	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BGL72	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BGK135	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BGK94	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60
BG10	TGACGGCCCTCTCGCGGGCGATCGTCTCCTTCTCGATGCGAGCGGAGATGCGGGCGAAGC	60

BG79	CCTCTGCGACCTTCTCGTTGAGCTGGTCGT-ACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
BG1032	CCTCTGCGACTTTCTCGTTGAGCTGGTCAT-ACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BG254	CCTCTGCGACTTTCTCGTTGAGCTGGTCGT-ACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
JX898210	CCTCTGCGACCTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
HM165226	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
AY072728	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
DQ090527	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
DQ090522	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
AY072727	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
HM165213	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
HM165208	CCTCTGCGACTTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
HM165216	CCTCTGCGACTTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BG257	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BG42	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BG15	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
HM165214	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
AY072725	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
JQ303245	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
X85958	CCTCTGCGACCTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
AY072723	CCTCTGCGACCTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
DQ090523	CCTCTGCGACTTTCTCGTTGAGCTGGTC-ATACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BGM13	CCTCTGCGACCTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
BGL78	CCTCTGCGACCTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
BGK10	CCTCTGCGACCTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
BGK111	CCTCTGCGACCTTCTCGTTGAGCTGGTCG-TACATCTTCTTCCTTTCTGCGTTCTCCGTG	119
BG411	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BGL63	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BGK110	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BGM67	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BGL72	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BGK135	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BGK94	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119
BG10	CCTCTGCGACTTTCTCGTTGAGCTGGTCA-TACATCTTCTTCCTCTCGGC GTTCTCCGTG	119

BG79	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
BG1032	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BG254	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGATCGTGTCTG	239
JX898210	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
HM165226	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
AY072728	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
DQ090527	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
DQ090522	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
AY072727	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
HM165213	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
HM165208	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTATCG	239
HM165216	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTATCG	239
BG257	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BG42	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BG15	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGTGTGTCTG	239
HM165214	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
AY072725	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
JQ303245	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
X85958	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
AY072723	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
DQ090523	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BGM13	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
BGL78	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
BGK10	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
BGK111	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGCGTGTCTG	239
BG411	GCGTCGTTATGGATGGCGATCTGGTTCTGGAGATTTGTCTCAACGTTGTTGAGTGTGTCTG	239
BGL63	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BGK110	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BGM67	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BGL72	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BGK135	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
BGK94	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGTGTGTCTG	239
BG10	GCGTCGTTGTGGATGGCGATCTGGTTCTGGAGGTTTCGTCTCGACGTTGTTGAGCGTGTCTG	239
	*****_*****_**_*****_*****_*_***	

BG79	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
BG1032	CCCATCTCAGCGAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BG254	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
JX898210	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
HM165226	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
AY072728	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
DQ090527	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
DQ090522	CCCATCTCAGCGAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
AY072727	CCCATCTCAGCGAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
HM165213	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
HM165208	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
HM165216	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BG257	CCCATCTCAGCGAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BG42	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BG15	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
HM165214	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
AY072725	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
JQ303245	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
X85958	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
AY072723	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
DQ090523	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGM13	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGL78	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGK10	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGK111	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
BG411	CCCATCTCCGCAAGGGA	CTTGC	GGAAGTTTGCAGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGL63	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGK110	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGM67	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGL72	CCCATCTCAGCGAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGK135	CCCATCTCAGCGAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BGK94	CCCATCTCAGCAAGAGAG	CTTGC	GGAAGTTTGCAGGCCATGGTGT	CGATCTCCTCCTTGATC	299
BG10	CCCATCTCAGCGAGAGAG	CTTGC	GGAAGTTTGT	-----	272
	***** _ * _ * _ *****				

BG79	GTTAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
BG1032	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BG254	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
JX898210	GTTAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
HM165226	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGAATGTCCTCGAACGAGGCC	359
AY072728	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
DQ090527	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
DQ090522	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
AY072727	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
HM165213	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
HM165208	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
HM165216	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BG257	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BG42	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BG15	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
HM165214	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
AY072725	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
JQ303245	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
X85958	GTTAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
AY072723	GTTAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
DQ090523	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGM13	GTTAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGL78	GTTAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGK10	GTCAAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGK111	GTCAAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
BG411	GTTAGGTACATGTTGTCGGCGGACTTCTTGACCTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGL63	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGK110	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGM67	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGCCCTCGAACGAGGCC	359
BGL72	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGK135	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BGK94	GTCAAGGTACATGTTGTCGGCAGACTTCTTGACTTCCTCGCGGATGTCCTCGAACGAGGCC	359
BG10	-----	272

BG79	TGGCGCTTCCTCGACTCCGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	413
BG1032	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
BG254	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
JX898210	TGGCGCTTCCTCGACTCCGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	413
HM165226	TGGCGCTTCCTCGACTCCGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
AY072728	TGGCGCTTCCTCGACTCCGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
DQ090527	TGGCGCTTCCTCGACTCCGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
DQ090522	TGGCGCTTCCTCGACTCCGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
AY072727	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
HM165213	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
HM165208	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGTGCGATGGCGTCC	413
HM165216	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
BG257	TGGGGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
BG42	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
BG15	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
HM165214	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
AY072725	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	413
JQ303245	TGGCGCTTCCTCGACTCGGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
X85958	TGGCGCTTCCTCGACTCCGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	413
AY072723	TGGCGCTTCCTCGACTCCGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	413
DQ090523	TGGCGCTTCCTCGACTCTGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	413
BGM13	TGGCGCTTCCTCGACT-CGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	412
BGL78	TGGCGCTTCCTCGACT-CGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	412
BGK10	TGGCGCTTCCTCGACT-CGTCTG-GATGAGCCTGTCGAGGTGCGCGATGGCGTCC	412
BGK111	TGGCGCTTCCTCGACT-CGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	412
BG411	TGGCGCTTCCTCGACTCCGTCTG-GATGAGCCTGTCGAGGTGTGCGATGGCGTCC	413
BGL63	TGGCGCTTCCTCGACT-TGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	412
BGK110	TGGCGCTTCCTCGACT-TGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	412
BGM67	TGGCGCTTCCTCGACT-TGTCTG-GATGAGTCTGTCAAGGTGCGCGATGGCGTCC	412
BGL72	TGGCGCTTCCTCGACT-TGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	412
BGK135	TGGCGCTTCCTCGAC-TTGTCTG-GATGAGTCTGCGAGGTGCGCGATGGCGTCC	412
BGK94	TGGCGCTTCCTCGA-CTTGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	412
BG10	-----CTCGACTCTGTCTG-GATGAGTCTGTGCGAGGTGCGCGATGGCGTCC	317

*****_**_*****_*****_**

Figure 4C. Multiple alignments of *gdh* sequences from this study with reference sequences obtained from GenBank, representing sub-assemblages of assemblage A and B.

CLUSTAL O(1.2.4) multiple sequence alignment

```

GDH84      TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
EU637582   CGACACTGACGTTCTGCCGGCGACATCGGCGTCGGCGCCCGCGAGATCGGCTATCTGTA 60
M84604     CGACACTGACGTTCTGCCGGCGACATCGGCGTCGGCGCCCGCGAGATCGGCTATCTGTA 60
GDM672     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTAACTGTT 60
GDL81      TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDK111     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
EU278608   CGACACTGACGTTCTGCCGGCGACATCGGCGTCGGCGCCCGCGAGATCGGGTACCTGTA 60
AY178738(isolate) TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDH371     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDM10      TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
AY178737   CGACACTGACGTTCTGCCGGCGACATCGGCGTCGGCGCCCGCGAGATCGGGTACCTGTA 60
GDH702     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDH530     CGACACTGACGTTCTGCCGGCGACATCGGCGTCGGCGCCCGCGAGATCGGGTACCTGTA 60
EF685684   TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDH727     TGACACCGANCTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
EU637587   TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDH738     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDH697     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDK135     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDK87      TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDL2       TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDK169     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
GDH591     TGACACCGACGTTCTGCTGGCGATATTGGCGTCGGCGGTCGCGAGATCGGTTATCTGTT 60
          *****

```


GDH84	TGGACAGTATAAGCGCTTCAGGAACGAGTTTACGGGCGTCCTCACGGGCAAGAACATCTT	120
EU637582	TGGGACGTACAAGCGCCTGAGGAACGAGTTACAGGCGTTCTCACAGGCAAGAACGTCAA	120
M84604	TGGGACGTACAAGCGCCTGAGGAACGAGTTACAGGCGTTCTCACAGGCAAGAACGTCAA	120
GDM672	TGGACAGTATAAGCGCCTCAGGAACGAATTCAGGGCGTCCTCACGGGCAAGAACATCAA	120
GDL81	TGGACAGTATAAGCGCCTCAGGAACGAATTCAGGGCGTCCTCACGGGCAAGAACATCAA	120
GDK111	TGGACAGTATAAGCGCCTCAGGAACGAGTTTACGGGCGTCCTCACGGGCAAGAACATCAA	120
EU278608	CGGACAGTACAAGCGCCTGAGGAACGAGTTACAGGCGTCCTCACAGGCAAGAACGTCAA	120
AY178738(isolate	TGGACAGTATAAGCGCCTCAGGAACGAGTTTACGGGCGTCCTCACGGGCAAGAACATCAA	120
GDH371	TGGACAGTATAAGCGCCTCAGGAACGAGTTTACGGGCGTCCTCACGGGCAAGAACATCAA	120
GDM10	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
AY178737	CGGACAGTACAAGCGCCTGAGGAACGAGTTACAGGCGTCCTCACAGGCAAGAACGTCAA	120
GDH702	TGGACAGTACAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDH530	CGGACAGTACAAGCGCCTGAGGAACGAGTTACAGGCGTCCTCACAGGCAAGAACGTCAA	120
EF685684	TGGACAGTACAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDH727	TGGACACTACAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
EU637587	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTTCTCACGGGCAAGAACATCAA	120
GDH738	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDH697	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDK135	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDK87	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDL2	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDK169	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
GDH591	TGGACAGTATAAGCGCCTCAGGAACGAGTTACAGGCGTCCTCACGGGCAAGAACATCAA	120
	**_* ** * ** : :	

GDH84	GTGGGGCGGGTCTCTTATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
EU637582	GTGGGGCGGGTCTCTTATCAGAGCCGAGAGGCCACGGGCTATGGTGCCGTCTACTTCCTGGA	180
M84604	GTGGGGCGGGTCTCTTATCAGAGCCGAGAGGCCACGGGCTATGGTGCCGTCTACTTCCTGGA	180
GDM672	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDL81	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDK111	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
EU278608	GTGGGGCGGGTCTCTTATCAGAGCCGAGAGGCCACGGGCTATGGCGCTGTCTACTTCCTGGA	180
AY178738(isolate	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDH371	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDM10	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCCGTCTACTTCCTGGA	180
AY178737	GTGGGGCGGGTCTCTTATCAGAGCCGAGAGGCCACGGGCTATGGCGCTGTCTACTTCCTGGA	180
GDH702	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDH530	GTGGGGCGGGTCTCTTATCAGAGCCGAGAGGCCACGGGCTATGGCGCTGTCTACTTCCTGGA	180
EF685684	GTGGGGCGGGTCTCTCATCAGAGCCAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDH727	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
EU637587	GTGGGGTGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDH738	GTGGGGCGGGTCTCTTATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDH697	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDK135	GTGGGGCGGGTCTCTCATCAAAACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDK87	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDL2	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDK169	GTGGGGCGGGTCTCTCATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180
GDH591	GTGGGGCGGGTCTCTTATCAGACCAAGAGGCCACAGGGTATGGAGCTGTCTACTTCCTGGA	180

GDH84	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
EU637582	GGAGATGTGCAAGGACAACAATACCGTGATCAGGGGCAAGAAATGTCCTCCTTTCTGGCTC	240
M84604	GGAGATGTGCAAGGACAACAATACCGTGATCAGGGGCAAGAAATGTCCTCCTTTCTGGCTC	240
GDM672	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDL81	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDK111	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
EU278608	GGAGATGTGCAAGGACAACAACACTGTGATCAGGAGTAAGAACGTCCTTCTTTCTGGCTC	240
AY178738(isolate	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDH371	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDM10	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
AY178737	GGAGATGTGCAAGGACAACAACACTGTGATCAGGGGTAAAGAACGTCCTCCTTTCTGGCTC	240
GDH702	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDH530	GGAGATGTGCAAGGACAACAACACTGTGATCAGGGGTAAAGAACGTCCTCCTTTCTGGCTC	240
EF685684	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTTCTCTCTGGCTC	240
GDH727	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
EU637587	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAAAAAGTCCTCCTCTCCGGCTC	240
GDH738	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDH697	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDK135	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDK87	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDL2	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDK169	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
GDH591	GGAGATGTGCAAGGATAACAACACCGTAATCAGGGGCAAGAACGTCCTCCTCTCTGGCTC	240
	***** ** *_*****_* *_* ***** * ** *****	

GDH84	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
EU637582	CGGCAACGTCGCTCAGTTTGTCTTGCAGAAAGCTCATTTAGCTCGGCGCAAAAGGTCCTCAC	300
M84604	CGGCAACGTCGCTCAGTTTGTCTTGCAGAAAGCTCATTTAGCTCGGCGCAAAAGGTCCTCAC	300
GDM672	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDL81	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDK111	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
EU278608	CGGCAACGTTGCCAGTTTGTCTTGCAGAAAGCTCATTTAGCTTGGCGCAAAAGGTCCTCAC	300
AY178738(isolate	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDH371	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDM10	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
AY178737	CGGCAACGTTGCCAGTTTGTCTTGCAGAAAGCTCATTTAGCTTGGCGCAAAAGGTCCTCAC	300
GDH702	TGGCAACGTTGCCAGTTTGTCTTGCAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDH530	CGGCAACGTTGCCAGTTTGTCTTGCAGAAAGCTCATTTAGCTTGGCGCAAAAGGTCCTCAC	300
EF685684	TGGCAACGTTGCCAGTTTGTCTTGCAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDH727	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
EU637587	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGCGCGAAGGTCCTCAC	300
GDH738	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDH697	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDK135	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDK87	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDL2	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDK169	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300
GDH591	TGGCAACGTTGCTCAGTACGCGTGCGAGAAAGCTCCTCCAGCTCGGTGCGAAGGTCCTCAC	300

***** _ * _ * _ * _ * _ * _ * _ * _ * _ * _ * _ * _ *

GDH84	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
EU637582	ATTCTCAGACTCCAAATGGGACCATTGTCGACA	332
M84604	ATTCTCAGACTCCAAATGGGACCATTGTCGACA	332
GDM672	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDL81	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDK111	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
EU278608	CTTCTCAGACTCCAAACGGGACCATTGTCGACA	332
AY178738(isolate	CTTCTCGGACTCCCAACGGGACTATCGTCGATA	332
GDH371	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDM10	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
AY178737	CTTCTCAGACTCCAAACGGGACCATTGTCGACA	332
GDH702	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDH530	CTTCTCAGACTCCAAACGGGACCATTGTCGACA	332
EF685684	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDH727	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
EU637587	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDH738	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDH697	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDK135	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDK87	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDL2	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDK169	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332
GDH591	CTTCTCGGACTCCAAACGGGACTATCGTCGATA	332

_ * _ * _ * _ * _ * _ * _ * _ * _ * _ * _ * _ *

Appendix 5. Total number of compounds occurring in each *Cryptosporidium* patient samples and non-infected samples.

Table 5A. Shows total number that each compound occurred in 11 *Cryptosporidium* samples and 11 control samples. It also shows their

Chi-Square test results and p-values for each compound between *Cryptosporidium* and each control. Highlighted compounds

show compounds which are biologically significant in the identification of Cryptosporidiosis.

No.	Compound name	RT	Freq in crypto	Freq in Neg	p-value
1	ethylalcohol	6.6	11	9	0.2
2	propanal	7.2	8	7	0.5
3	Acetone	7.4	6	7	0.5
4	isopropyl alcohol	7.6	10	9	0.5
5	carbon disulfide	7.8	0	2	0.2
6	acetic acid, methyl ester	8.0	2	1	0.5
7	acetaldehyde, hydroxy-	8.3	0	0	-
8	cyclopentane	8.3	7	1	0.01
9	2-methylpropanal	9.0	6	10	0.07
10	1-propanol	9.6	0	11	0.01
11	2,3-butanedione	10.3	10	10	0.8
12	acetic anhydride	10.4	1	1	0.8
13	2-butanone	10.6	1	0	0.5
14	ethyl acetate	10.6	3	5	0.3
15	2-methyl-1-propanal	12.0	3	2	0.5
16	butanal, 3-methyl-	12.5	3	3	0.7
17	hydrogen azide	12.6	0	1	0.5

18	acetic acid	12.6	10	11	0.5
19	acetic anhydride	12.7	1	0	0.5
20	ethyl-1-propenyl ether	12.8	1	1	0.8
21	butanal, 2-methyl-	12.9	6	6	0.7
22	2-ethylfuran	13.5	3	3	0.7
23	1-penten-3-one	13.9	2	3	0.5
24	2-pentanone	14.0	0	1	0.5
25	1-penten-3-ol	14.0	1	2	0.5
26	pentanal	14.2	7	8	0.5
27	acetic acid, methyl ester	14.4	1	0	0.5
28	3-hydroxy-2-butanone	16.1	10	5	0.03
29	1-butanol, 3-methyl-	16.2	0	1	0.5
30	propanoic acid	16.3	5	6	0.5
31	3-penten-2-one	16.4	1	1	0.8
32	Toluene	16.6	0	3	0.1
33	Octane	16.6	1	1	0.8
34	acetic acid, 2-methylpropyl ester	16.7	0	1	0.5
35	2-pentenal, (e)	17.1	0	2	0.2
36	1-pentanol	17.5	1	2	0.5
37	butanoic acid, ethyl ester	17.7	1	4	0.2
38	2-hexanone	18.1	1	0	0.5
39	Hexanal	18.3	10	11	0.5
40	propanoic acid, 2-methyl-	18.4	1	2	0.5
41	pentanoic acid	19.5	0	1	0.5
42	butanoic acid	19.6	6	7	0.5
43	ethoxyacetylene	19.6	1	0	0.5

44	acetic acid, [(1,1-dimethylethyl)thio]-	19.8	1	0	0.5
45	2,2-dimethyl-propanoic acid	19.9	2	0	0.2
46	ethylbenzene	20.3	2	2	0.7
47	1,3-cyclopentadiene, 5-(1-methylethylidene)-	20.6	2	1	0.5
48	1-butanol, 3-methyl-, acetate	20.8	0	1	0.5
49	3-methyl-1-butanol	20.8	1	0	0.5
50	3-furaldehyde	20.8	1	0	0.5
51	2-hexanal	21.1	1	0	0.5
52	4-heptanone	21.2	1	0	0.5
53	1-hexanol	21.4	1	0	0.5
54	butanoic acid, 3-methyl-	21.5	10	8	0.3
55	butanoic acid, 2-methyl-	21.8	7	7	0.7
56	Acetone	22.0	0	1	0.5
57	2-heptanone	22.0	4	3	0.5
58	acetic acid, pentyl ester	22.2	0	1	0.5
59	heptanal	22.3	8	8	0.7
60	Propane	22.3	1	0	0.5
61	pentanoic acid, 3-methyl-	23.0	0	2	0.2
62	cyclohexanone	23.1	0	1	0.5
63	cyclopentanone, 3-methyl-	23.1	0	1	0.5
64	benzene, 1-ethyl-2-methyl-/ acetophenone	24.2	1	1	0.8
65	1,3,5-triazine	24.8	1	0	0.5
66	furan, 2-methyl-	24.8	0	1	0.5
67	2-pentylfuran	24.8	9	7	0.3
68	2-heptanal	25.0	7	3	0.1
69	1-hepten-3-one	25.2	1	1	0.8

70	2,3-octanedione	25.2	1	0	0.5
71	2-methylpropanoic acid, anhydride	25.2	1	0	0.5
72	5-amino-2-methyl-2h-tetrazole	25.2	0	1	0.5
73	1-octen-3-ol	25.3	2	0	0.2
74	ethaneperoxoic acid, 1-cyano-1-phenylpentyl ester	25.5	1	0	0.5
75	dimethyl trisulfide	25.4	0	1	0.5
76	3-octanone	25.4	2	2	0.7
77	1,3,4-oxadiazole-2-thiol, 5-cyclopropyl-	25.5	0	1	0.5
78	1,1-difluoro-2-methyl-3-ethyl cyclopropane	25.5	1	0	0.5
79	5-hepten-2-one, 6-methyl-	25.6	3	4	0.5
80	ethanol, 2-(ethylthio)-	25.6	0	1	0.5
81	benzaldehyde	25.6	8	5	0.2
82	2-octanone	25.7	2	3	0.5
83	acetic acid, hexylester	25.8	0	1	0.5
84	2,2,4,4-tetramethyloctane	25.8	7	5	0.3
85	pyrazine, trimethyl-	25.9	4	3	0.5
86	Octanal	26.1	7	6	0.5
87	d-limonene	26.2	1	0	0.5
88	pentanoic acid	26.4	2	2	0.7
89	hexanoic acid	26.4	4	6	0.3
90	2-butanone, 3-methoxy-3-methyl-	26.4	1	0	0.5
91	1h-pyrrole, 2-methyl-	26.8	1	0	0.5
92	1h-tetrazole	27.2	0	1	0.5
93	butanoic acid, 3-methylbutyl ester / pyrrolidine	27.2	1	1	0.8
94	2,4-pentadienoic acid, 1-cyclopenten-3-on-1-yl ester	27.3	1	0	0.5
95	2(3h)-furanone	27.6	0	1	0.5

96	5-ethylcyclopent-1-enecarboxaldehyde	27.8	1	1	0.8
97	ethanone, 2-cyclohexyl-1-(1-methyl-1h-imidazol-4-yl)-	27.8	1	0	0.5
98	3-octen-2-one	27.8	0	2	0.2
99	Phenol	28.1	4	3	0.5
100	2-octenal, (e)	28.6	3	3	0.7
101	propanenitrile, 3-[(phenylmethylamino)-	28.7	1	0	0.5
102	benzeneacetaldehyde	28.7	6	0	0.5
103	pyrazine, tetramethyl-	28.7	1	5	0.5
104	acetic acid ethenyl ester	28.8	0	0	-
105	3,5-octadien-2-one	29.1	2	4	0.3
106	2-nonanone	29.2	5	2	0.2
107	acetophenone	29.5	1	0	0.5
108	Nonanal	29.5	10	6	0.1
109	2h-1-benzopyran, 3,4-dihydro-	30.0	0	1	0.5
110	benzene, 1-methyl-2-(1-methylethyl)-	30.0	2	0	0.2
111	5-ethyldihydro-2(3h)-furanone	30.4	0	2	0.2
112	2,6-nonadienal, (e,e)-	31.9	0	1	0.5
113	2-nonenal, (e)-	32.0	2	4	0.3
114	octanoic acid	32.7	0	2	0.2
115	Decanal	32.8	0	2	0.2
116	benzene, 1,3-bis(1,1-dimethylethyl)-	33.7	2	1	0.5
117	2-undecanone	35.5	1	0	0.5
118	benzenemethanol, à,4-dimethyl-	37.2	0	1	0.5

Appendix 6. A selection of certificates from some of my presentations, training courses and workshops



IV INTERNATIONAL **GIARDIA**
& **cryptosporidium**
CONFERENCE | 31 JANUARY – 3 FEBRUARY 2012
WELLINGTON, NEW ZEALAND

Certificate of Attendance

Chadaporn Nuchjangreed

attended the

**IV International Giardia & Cryptosporidium Pre-Conference
Workshop**

at

Massey University, Wellington

on

30 January 2012

Elaine Moriarty
Organising Committee : IV International
Giardia and Cryptosporidium Pre-
Conference Workshop 2012



IV INTERNATIONAL **GIARDIA**
& **cryptosporidium**
CONFERENCE | 31 JANUARY - 3 FEBRUARY 2012
WELLINGTON, NEW ZEALAND

Certificate of Appreciation

Chadaporn Nuchjangreed

for attending

IV International Giardia & Cryptosporidium Conference

at

Te Papa, Wellington

from

31 January to 3 February 2012



Alex Grinberg
Chair, IV International Giardia &
Cryptosporidium Conference 2012



Certificate of Appreciation

This is presented to

Chadaporn Nuchjangreed

for contribution as a selected abstract presenter in
The 3rd Health Challenge Thailand Academic Conference
June 15th, 2013

Office of Educational Affairs, Royal Thai Embassy, London, UK

A handwritten signature in black ink, likely belonging to Dr. Piyawat Sivaraks.

Dr. Piyawat Sivaraks
Minister (for Education)
Office of Educational Affairs, UK

A handwritten signature in black ink, likely belonging to Dr. Usa Kullaprawithaya.

Dr. Usa Kullaprawithaya
Minister Counsellor (Science and Technology)
Office of Science and Technology, Brussels



CERTIFICATE OF TRAINING

This certificate is awarded to:

Chadaporn Nuchjangreed


Who has trained in the 40-hours

MOLECULAR PARASITOLOGY WORKSHOP

August 7th – 10th of 2014 at CINVESTAV, Zacatenco

Mexico city


Juan David Ramirez, Ph.D.
Universidad del Rosario


Rebeca Manning Cella, Ph.D.
CINVESTAV


Fidel de la Cruz Hernández-Hernández, Ph.D.
CINVESTAV



The Mexican Society of Parasitology and the World Federation of Parasitologists



Acknowledge:

An Investigation of Giardiasis and Cryptosporidiosis in Cambodia and Malawi

Chadaporn Nuchjangreed

Nadine Randle , Dove Winifred , Nigel Cunliffe , Catrin Moore , Christopher Parry , Nicholas Beeching , Wendi
Bailey , Varun Kumar , Caroline Corless , Rachel Chalmers , Jonathan Wastling

For Presenting a free oral paper in the

**13th INTERNATIONAL
CONGRESS OF PARASITOLOGY**
AUGUST 10th - 15th, 2014
at the Camino Real Hotel
Mexico City

Dra. Ana Filser
President
Mexican Society of Parasitology

Dr. Alan Cowman
President
World Federation of Parasitologists



Certificate of Participation

This is presented to

Chadaporn Nuchjangreed

For an abstract submission to

The 4th Health Challenge Thailand Academic Conference

September 6th, 2014

Office of Educational Affairs, Royal Thai Embassy, London, UK

Miss Orrawan Nuypakdee
Minister (Education)
Office of Education Affairs, UK

Dr Usa Kullaprawithaya
Minister Counselor (Science and Technology)
Office of Science and Technology, Brussels

REFERENCES

- Abboud P, Lemée V, Gargala G, et al.** (2001). Successful treatment of metronidazole-and albendazole-resistant giardiasis with nitazoxanide in a patient with acquired immunodeficiency syndrome. *Clin Infect Dis*, **32**, 1792.
- Abd El Kader, N. M., Blanco, M. A., Ali-Tammam, M., Abd El Ghaffar Ael, R., Osman, A., El Sheikh, N., Rubio, J. M. and de Fuentes, I.** (2012). Detection of *Cryptosporidium parvum* and *Cryptosporidium hominis* in human patients in Cairo, Egypt. *Parasitol Res*, **110**, 161-166.
- Abdel-Hafeez, E. H., Ahmad, A. K., Ali, B. A. and Moslam, F. A.** (2012). Opportunistic parasites among immunosuppressed children in Minia District, Egypt. *Korean J Parasitol*, **50**, 57-62.
- Abe, N., Matsubayashi, M., Kimata, I. and Iseki, M.** (2006). Subgenotype analysis of *Cryptosporidium parvum* isolates from humans and animals in Japan using the 60-kDa glycoprotein gene sequences. *Parasitol Res*, **99**, 303-305.
- Abu SN, Thompson PN, Jori F, Frean J, Poonsamy B, duPlessis D, et al.** (2013). Genetic characterization of *Cryptosporidium* spp. in diarrhoeic children from four provinces in South Africa. *Zoonoses Public Health*, **60**, 154-9.
- Adam, R. D.** (2001). Biology of *Giardia lamblia*. *Clin Microbiol Rev*, **14**, 447-475.
- Adamu, H., Petros, B., Hailu, A. and Petry, F.** (2010). Molecular characterization of *Cryptosporidium* isolates from humans in Ethiopia. *Acta Trop*, **115**, 77-83.
- Adamu, H., Wegayehu, T. and Petros, B.** (2013). High prevalence of diarrhoeagenic intestinal parasite infections among non-ART HIV patients in Fitch Hospital, Ethiopia. *PLoS One*, **8**, e72634. doi: 10.1371/journal.pone.0072634.

- Ajjampur, S. S., Gladstone, B. P., Selvapandian, D., Muliyl, J. P., Ward, H. and Kang, G. (2007).** Molecular and spatial epidemiology of Cryptosporidiosis in children in a semiurban community in South India. *J Clin Microbiol*, **45**, 915-920.
- Ajjampur, S. S., Sankaran, P., Kannan, A., Sathyakumar, K., Sarkar, R., Gladstone, B. P. and Kang, G. (2009).** *Giardia duodenalis* assemblages associated with diarrhea in children in South India identified by PCR-RFLP. *Am J Trop Med Hyg*, **80**, 16-19.
- Ajjampur, S. S., Liakath FB, Kannan A, Rajendran P, Sarkar R, Moses PD, Simon A, Agarwall, Mathew A, O'Connor R, Ward H, Kang G. (2010).** Multisite study of cryptosporidiosis in children with diarrhea in India. *J Clin Microbiol* **48**, 2075-2081.
- Akiyoshi, D. E., Tumwine, J. K., Bakeera-Kitaka, S. and Tzipori, S. (2006).** Subtype analysis of *Cryptosporidium* isolates from children in Uganda. *J Parasitol*, **92**, 1097-1100.
- Aldeen, W. E., Carroll, K., Robison, A., Morrison, M. and Hale, D. (1998).** Comparison of nine commercially available enzyme-linked immunosorbent assays for detection of *Giardia lamblia* in fecal specimens. *J Clin Microbiol*, **36**, 1338-1340.
- Allen MP, Blake JF, Bryce DK, Haggan ME, Liras S, McLean S, Segelstein BE. (2000).** Design, synthesis and biological evaluation of 3-amino-3-phenylpropionamide derivatives as novel mu opioid receptor ligands. *Bioorg Med Chem Lett*. **10**, 523-6.
- Almeida, A., Pozio, E. and Caccio, S. M. (2010).** Genotyping of *Giardia duodenalis* Cysts by New Real-Time PCR Assays for Detection of Mixed Infections in Human Samples. *Appl Environ Microbiol*, **76**, 1895-1901.
- Alves, M., Ribeiro, A. M., Neto, C., Ferreira, E., Benoliel, M. J., Antunes, F. and Matos, O. (2006).** Distribution of *Cryptosporidium* species and subtypes in water samples in Portugal: a preliminary study. *J Eukaryot Microbiol*, **53 Suppl 1**, S24-25.

- Alves, M., Xiao, L., Sulaiman, I., Lal, A. A., Matos, O. and Antunes, F.** (2003). Subgenotype analysis of *Cryptosporidium* isolates from humans, cattle, and zoo ruminants in Portugal. *J Clin Microbiol*, **41**, 2744-2747.
- Amadi, B., Kelly, P., Mwiya, M., Mulwazi, E., Sianongo, S., Changwe, F., Thomson, M., Hachungula, J., Watuka, A., Walker-Smith, J. and Chintu, C.** (2001). Intestinal and systemic infection, HIV, and mortality in Zambian children with persistent diarrhea and malnutrition. *J Pediatr Gastroenterol Nutr*, **32**, 550-554.
- Amar, C. F. L., Dear, P. H., Pedraza-Díaz, S., Looker, N., Linnane, E. and McLauchlin, J.** (2002). Sensitive PCR-restriction fragment length polymorphism assay for detection and genotyping of *Giardia duodenalis* in human feces. *J Clin Microbiol*, **40**, 446-452.
- American Society of Health-System Pharmacists.** Azithromycin Retrieved Aug 1, 2015.
- American Society of Health-System Pharmacists.** Clarithromycin. Retrieved September 4, 2015.
- American Society of Health-System Pharmacists.** Rifabutin Retrieved 8 December 2016.
- Ankarklev, J., Hestvik, E., Lebbad, M., Lindh, J., Kaddu-Mulindwa, D. H., Andersson, J. O., Tylleskär, T., Tumwine, J. K. and Svärd, S. G.** (2012). Common Coinfections of *Giardia intestinalis* and *Helicobacter pylori* in Non-Symptomatic Ugandan Children. *PLoS Negl Trop Dis*, **6**.
- Ankarklev J, Svärd SG, Lebbad M.** Allelic sequence heterozygosity in single *Giardia* parasites. *BMC Microbiol*. 2012;12:65.
- Appelbee, A. J., Thompson, R. C. and Olson, M. E.** (2005). *Giardia* and *Cryptosporidium* in mammalian wildlife-current status and future needs. *Trends Parasitol*, **21**, 370-376.
- Araujo, A. J., Kanamura, H. Y., Almeida, M. E., Gomes, A. H., Pinto, T. H. and Da Silva, A. J.** (2008). Genotypic identification of *Cryptosporidium* spp. isolated from HIV-infected patients and immunocompetent children of Sao Paulo, Brazil. *Rev Inst Med Trop Sao Paulo*, **50**, 139-143.

- Areeshi, M., Dove, W., Papaventsis, D., Gatei, W., Combe, P., Grosjean, P., Leatherbarrow, H. and Hart, C. A.** (2008). *Cryptosporidium* species causing acute diarrhoea in children in Antananarivo, Madagascar. *Ann Trop Med Parasitol*, **102**, 309-315.
- Arthur C. L. and Pawliszyn J.** (1990). Solid Phase Microextraction with Thermal Desorption using Fused Silica Optical Fibres. *Anal. Chem.* **62**, 2145.
- Atta-Owusu C.** (2008). The prevalence of intestinal parasites among children at selected Day Care Centres in Accra, Ghana. Accra, Ghana University of Ghana.
- Bajer, A., Bednarska, M., Caccio, S. M., Wolska-Kusnierz, B., Heropolitanska-Pliszka, E., Bernatowska, E., Wielopolska, M., Paziewska, A., Welc-faleciak, R. and Sinski, E.** (2008). Genotyping of *Cryptosporidium* isolates from human clinical cases in Poland. *Parasitol Res*, **103**, 37-42.
- Barr, S. C., Bowman, D. D., Frongillo, M. F. and Joseph, S. L.** (1998). Efficacy of a drug combination of praziquantel, pyrantel pamoate, and febantel against giardiasis in dogs. *Am J Vet Res*, **59**, 1134-1136.
- BBC.** Boil water warning 'precaution' 2 September 2008. Retrieved 7 September 2009.
- Behera, B., Mirdha, B. R., Makharia, G. K., Bhatnagar, S., Dattagupta, S. and Samantaray, J. C.** (2008). Parasites in patients with malabsorption syndrome: a clinical study in children and adults. *Dig Dis Sci*, **53**, 672-679.
- Bern, C., Hernandez, B., Lopez, M. B., Arrowood, M. J., De Merida, A. M. and Klein, R. E.** (2000). The contrasting epidemiology of *Cyclospora* and *Cryptosporidium* among outpatients in Guatemala. *Am J Trop Med Hyg*, **63**, 231-235.
- Bern, C., Ortega, Y., Checkley, W., Roberts, J. M., Lescano, A. G., Cabrera, L., Verastegui, M., Black, R. E., Sterling, C. and Gilman, R. H.** (2002). Epidemiologic differences between cyclosporiasis and cryptosporidiosis in Peruvian children. *Emerg Infect Dis*, **8**, 581-585.

- Bertrand, I., Albertini, L. and Schwartzbrod, J.** (2005). Comparison of two target genes for detection and genotyping of *Giardia lamblia* in human feces by PCR and PCR-restriction fragment length polymorphism. *J Clin Microbiol*, **43**, 5940-5944.
- Betancourt, WQ; Rose, JB** (2004). Drinking water treatment processes for removal of *Cryptosporidium* and *Giardia*. *Vet Parasitol*. **126** (1-2), 219-34.
- Bhattacharya, S. and Mandal, S. S.** (1997). Interaction of surfactants with DNA. Role of hydrophobicity and surface charge on intercalation and DNA melting. *Biochim Biophys Acta*, **1323**, 29-44.
- Blackburn, B. G., Mazurek, J. M., Hlavsa, M., Park, J., Tillapaw, M., Parrish, M., Salehi, E., Franks, W., Koch, E., Smith, F., Xiao, L., Arrowood, M., Hill, V., da Silva, A., Johnston, S. and Jones, J. L.** (2006). Cryptosporidiosis associated with ozonated apple cider. *Emerg Infect Dis*, **12**, 684-686.
- Boaitey, Y. A., Nkrumah, B., Idriss, A. and Tay, S. C.** (2012). Gastrointestinal and urinary tract pathogenic infections among HIV seropositive patients at the Komfo Anokye Teaching Hospital in Ghana. *BMC Res Notes*, **5**, 454.
- Boontanom, P., Mungthin, M., Tan-ariya, P., Naaglor, T. and Leelayoova, S.** (2011). Epidemiology of giardiasis and genotypic characterization of *Giardia duodenalis* in preschool children of a rural community, central Thailand. *Trop Biomed*, **28**, 32-39.
- Brayfield, A, ed.** (2014) (14 January 2014). Metronidazole. Martindale: The Complete Drug Reference. *Pharmaceutical Press*. Retrieved 3 April 2014.
- Breathnach, A. S., McHugh, T. D. and Butcher, P. D.** (2010). Prevalence and clinical correlations of genetic subtypes of *Giardia lamblia* in an urban setting. *Epidemiol Infect*, **138**, 1459-1467.
- Broglia A, Weitzel T, Harms G, Cacciò SM, Nöckler K.** (2013). Molecular typing of *Giardia duodenalis* isolates from German travellers. *Parasitol Res*, **112**, 3449-56.

- Brooks, GF, Butel, Janet S, Morse, Stephen A.** (2004). Jawetz, Melnick, & Adelberg's Medical Microbiology (23rd ed.). *New York: Lange Medical Books/McGraw Hill*, 684-5.
- Buret A, Hardin JA, Olson ME, Gall DG.** (1992). Pathophysiology of small intestinal malabsorption in gerbils infected with *Giardia lamblia*. *Gastroenterol*, **103**, 506-513.
- Buret A, Gall DG, Olson ME.** (1991). Growth, activities of enzymes in the small intestine, and ultrastructure of microvillous border in gerbils infected with *Giardia duodenalis*. *Parasitol Res*, **77**, 109-114.
- Burton, A. J., Nydam, D. V., Dearen, T. K., Mitchell, K., Bowman, D. D. and Xiao, L.** (2010). The prevalence of *Cryptosporidium*, and identification of the *Cryptosporidium* horse genotype in foals in New York State. *Vet Parasitol*, **174**, 139-144.
- Busch, K. L.** (2010). Mass spectral libraries: The next generation. *Spectroscopy*. **25**.
- Bushen, O. Y., Kohli, A., Pinkerton, R. C., Dupnik, K., Newman, R. D., Sears, C. L., Fayer, Cacciò, S. M., Beck, R., Lalle, M., Marinculic, A. and Pozio, E.** (2008). Multilocus genotyping of *Giardia duodenalis* reveals striking differences between assemblages A and B. *Int J Parasitol*, **38**, 1523-1531.
- Cabada MM, White AC, Venugopalan P, Sureshbabu J.** (2015). Bronze MS, ed. Cryptosporidiosis Treatment & Management. Medscape. Web MD. Retrieved 8 January 2016.
- Cacciò, S. M., De Giacomo, M. and Pozio, E.** (2002). Sequence analysis of the β -giardin gene and development of a polymerase chain reaction-restriction fragment length polymorphism assay to genotype *Giardia duodenalis* cysts from human faecal samples. *Int J Parasitol*, **32**, 1023-1030.
- Caccio SM et al.** (2005). Unravelling *Cryptosporidium* and *Giardia* epidemiology. *Trends Parasitol*, **21**, 430-437.
- Caccio SM, Ryan U.** 2008. Molecular epidemiology of giardiasis. *Mol Biochem Parasitol*, **160**, 75-80.

- Cacopardo B, Patamia I, Bonaccorso V, et al.** (1995). Synergic effect of albendazole plus metronidazole association in the treatment of metronidazole-resistant giardiasis. *Clin Ter*, **146**, 761.
- Calderaro, A., Gorrini, C., Montecchini, S., Peruzzi, S., Piccolo, G., Rossi, S., Gargiulo, F., Manca, N., Dettori, G. and Chezzi, C.** (2010). Evaluation of a real-time polymerase chain reaction assay for the laboratory diagnosis of giardiasis. *Diagn Microbiol Infect Dis*, **66**, 261-267.
- Cama, V., Gilman, R. H., Vivar, A., Ticona, E., Ortega, Y., Bern, C. and Xiao, L.** (2006). Mixed *Cryptosporidium* infections and HIV. *Emerg Infect Dis*, **12**, 1025-1028.
- Cama, V. A., Bern, C., Roberts, J., Cabrera, L., Sterling, C. R., Ortega, Y., Gilman, R. H. and Xiao, L.** (2008). *Cryptosporidium* species and subtypes and clinical manifestations in children, Peru. *Emerg Infect Dis*, **14**, 1567-1574.
- Cama, V. A., Bern, C., Sulaiman, I. M., Gilman, R. H., Ticona, E., Vivar, A., Kawai, V., Vargas, D., Zhou, L. and Xiao, L.** (2003). *Cryptosporidium* species and genotypes in HIV-positive patients in Lima, Peru. *J Eukaryot Microbiol*, **50 Suppl**, 531-533.
- Cama, V. A., Ross, J. M., Crawford, S., Kawai, V., Chavez-Valdez, R., Vargas, D., Vivar, A., Ticona, E., Navincopa, M., Williamson, J., Ortega, Y., Gilman, R. H., Bern, C. and Xiao, L.** (2007). Differences in clinical manifestations among *Cryptosporidium* species and subtypes in HIV-infected persons. *J Infect Dis*, **196**, 684-691.
- Centers for Disease Control and Prevention.** Cryptosporidiosis 5 February 2009.
- Centers for Disease Control and Prevention.** CDC. Parasites - Giardia, Prevention & Control". Retrieved 26 April 2015.
- Cerdan-Calero, M., Sendra, J. M. and Sentandreu, E.** (2012). Gas chromatography coupled to mass spectrometry analysis of volatiles, sugars, organic acids and aminoacids in Valencia Late orange juice and reliability of the Automated Mass Spectral Deconvolution and

- Identification System for their automatic identification and quantification. *J Chromatogr A*, **1241**, 84-95.
- Chalmers, R. M., Robinson, G., Elwin, K., Hadfield, S. J., Xiao, L., Ryan, U., Modha, D. and Mallaghan, C.** (2009). *Cryptosporidium* sp. rabbit genotype, a newly identified human pathogen. *Emerg Infect Dis*, **15**, 829-830.
- Checkley W, Epstein LD, Gilman RH, et al.** (1998). Effects of *Cryptosporidium parvum* infection in Peruvian children: growth faltering and subsequent catch-up growth. *Am J Epidemiol*; **148**, 497.
- Chen W, Harp JA, Harmsen AG.** (2003). *Cryptosporidium parvum* infection in gene-targeted B cell-deficient mice. *J. Parasitol*, **89** (2), 391-3.
- Chesnaye, N., Sinuon, M., Socheat, D., Koporc, K. and Mathieu, E.** (2011). Treatment coverage survey after a school-based mass distribution of mebendazole: Kampot Province, Cambodia. *Acta Trop*, **118**, 21-26.
- Cranendonk, R. J., Kodde, C. J., Chipeta, D., Zijlstra, E. E. and Sluiter, J. F.** (2003). *Cryptosporidium parvum* and *Isospora belli* infections among patients with and without diarrhoea. *East Afr Med J*, **80**, 398-401.
- Cunliffe, N., Wittel, D. and Ngwira, B.** (2009a). History of *rotavirus* research in children in Malawi: the pursuit of a killer. *Malawi Med J*, **21**, 113-115.
- Cunliffe, N. A., Ngwira, B. M., Dove, W., Nakagomi, O., Nakagomi, T., Perez, A., Hart, C. A., Kazembe, P. N. and Mwansambo, C. C.** (2009b). Serotype g12 *rotaviruses*, Lilongwe, Malawi. *Emerg Infect Dis*, **15**, 87-90.
- Dagci H, Ustun S, Taner MS, et al.** (2002). Protozoan infections and intestinal permeability. *Acta Trop*, **81**, 1-5.

- Das, P., Roy, S. S., MitraDhar, K., Dutta, P., Bhattacharya, M. K., Sen, A., Ganguly, S., Bhattacharya, S. K., Lal, A. A. and Xiao, L.** (2006). Molecular characterization of *Cryptosporidium* spp. from children in Kolkata, India. *J Clin Microbiol*, **44**, 4246-4249.
- Davidson RN, den Boer M, Ritmeijer K.** (2009). Paromomycin. *Trans R Soc Trop Med Hyg*, **103**(7), 653-660.
- De Lacy Costello B., Ewer A. K., Garner C. E., Probert C. S. J., Ratcliffe N. M., Smith S.** (2008). An analysis of volatiles in the headspace of the faeces of neonates. *J Breath Res*, **2**(3).
- Di Cristanziano, V., Santoro, M., Parisi, F., Albonico, M., Shaali, M. A., Di Cave, D. and Berrilli, F.** (2014). Genetic characterization of *Giardia duodenalis* by sequence analysis in humans and animals in Pemba Island, Tanzania. *Parasitol Int*, **63**, 438-441.
- Dixon, B., Parrington, L., Cook, A., Pintar, K., Pollari, F., Kelton, D. and Farber, J.** (2011). The potential for zoonotic transmission of *Giardia duodenalis* and *Cryptosporidium* spp. from beef and dairy cattle in Ontario, Canada. *Vet Parasitol*, **175**, 20-26.
- Dorny, P., Stoliaroff, V., Charlier, J., Meas, S., Sorn, S., Chea, B., Holl, D., Van Aken, D. and Vercruysse, J.** (2011). Infections with gastrointestinal nematodes, *Fasciola* and *Paramphistomum* in cattle in Cambodia and their association with morbidity parameters. *Vet Parasitol*, **175**, 293-299.
- DrugBank**, ed. (22 March 2017). "Rifaximin" – via DrugBank.ca.
- DuPont HL, Chappell CL, Sterling CR, et al.** (1995). The infectivity of *Cryptosporidium parvum* in healthy volunteers. *N Engl J Med*, **332**, 855.
- Dutta AK, Phadke MA, Bagade AC, et al.** (1994). A randomised multicentre study to compare the safety and efficacy of albendazole and metronidazole in the treatment of giardiasis in children. *Indian J Pediatr*, **61**, 689.

- Edwards, David I.** (1993). Nitroimidazole drugs - action and resistance mechanisms. I. Mechanism of action. *Journal of Antimicrobial Chemotherapy*, **31**, 9-20.
- Edison, NJ.** (2012). Heritage Pharmaceuticals; Paromycin sulfate capsules (package insert).
- El-Shazly AM, Mowafy N, Soliman M, El-Bendary M, Morsy AT, Ramadan NI, et al.** (2004). Egyptian genotyping of *Giardia lamblia*. *J Egypt Soc Parasitol*, **34**, 265-80.
- Enriquez, C., Nwachuku, N. and Gerba, C. P.** (2001). Direct exposure to animal enteric pathogens. *Rev Environ Health*, **16**, 117-131.
- Enweronu-Laryea, C. C., Sagoe, K. W., Mwenda, J. M. and Armah, G. E.** (2014). Severe acute *rotavirus* gastroenteritis in children less than 5 years in southern Ghana: 2006-2011. *Pediatr Infect Dis J*, **33 Suppl 1**, S9-S13.
- Exner, M; Gornik, V** (2004). Parasitic zoonoses transmitted by drinking water. Giardiasis and Cryptosporidiosis. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*. **47** (7): 698-704.
- Fahmy HM, El-Serougi AO, El Deeb HK, Hussein HM, Abou-Seri HM, Klotz C, Aebischer T, Khalifa ES, Mohamed K.** (2015). *Giardia duodenalis* assemblages in Egyptian children with diarrhea. *Eur J Clin Microbiol Infect Dis*. **34**(8), 1573-1581.
- Färnert A, Lindberg J, Gil P, et al.** (2003). Evidence of *Plasmodium falciparum* malaria resistant to atovaquone and proguanil hydrochloride: case reports". *BMJ*. **326** (7390), 628-9.
- Farthing, M., Feldman, R., Finch, R., Fox, R., Leen, C., Mandal, B., Moss, P., Nathwani, D., Nye, F., Percival, A., Read, R., Ritchie, L., Todd, W. T. and Wood, M.** (1996). The management of infective gastroenteritis in adults. A consensus statement by an expert panel convened by the British Society for the Study of Infection. *J Infect*, **33**, 143-152.
- Farthing, M. J.** (2000). Diarrhoea: a significant worldwide problem. *Int J Antimicrob Agents*, **14**, 65-69.

- Faubert G.** (2000). Immune response to *Giardia duodenalis*. *Clin Microbiol Rev*, **13**, 35-54.
- Fayer, R., Morgan, U. and Upton, S. J.** (2000a). Epidemiology of *Cryptosporidium*: transmission, detection and identification. *Int J Parasitol*, **30**, 1305-1322.
- Fayer, R., Trout, J. M., Walsh, E. and Cole, R.** (2000b). Rotifers ingest oocysts of *Cryptosporidium parvum*. *J Eukaryot Microbiol*, **47**, 161-163.
- Felsenstein J.** (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Syst Biol*, **27**(4), 401-410.
- Felsenstein J.** (1985). Phylogenies and the Comparative Method, *The American Naturalist*, **125**, 1-15.
- Feltus, D. C., Giddings, C. W., Schneck, B. L., Monson, T., Warshauer, D. and McEvoy, J. M.** (2006). Evidence supporting zoonotic transmission of *Cryptosporidium* spp. in Wisconsin. *J Clin Microbiol*, **44**, 4303-4308.
- Feng, X., Akiyoshi, D. E., Widmer, G. and Tzipori, S.** (2007a). Characterisation of subtilase protease in *Cryptosporidium parvum* and *C. hominis*. *J Parasitol*, **93**, 619-626.
- Feng, Y., Alderisio, K. A., Yang, W., Blancero, L. A., Kuhne, W. G., Nadareski, C. A., Reid, M. and Xiao, L.** (2007b). *Cryptosporidium* genotypes in wildlife from a new york watershed. *Appl Environ Microbiol*, **73**, 6475-6483.
- Feng Y. Alderisio K.A. Yang W. et al.** (2007a) *Cryptosporidium* genotypes in wildlife from a New York watershed. *Appl Environ Microbiol*, **73**, 6475-6483.
- Feng, Y., Li, N., Duan, L. and Xiao, L.** (2009). *Cryptosporidium* genotype and subtype distribution in raw wastewater in Shanghai, China: evidence for possible unique *Cryptosporidium hominis* transmission. *J Clin Microbiol*, **47**, 153-157.

- Feng, Y., Ortega, Y., He, G., Das, P., Xu, M., Zhang, X., Fayer, R., Gatei, W., Cama, V. and Xiao, L.** (2007c). Wide geographic distribution of *Cryptosporidium bovis* and the deer-like genotype in bovines. *Vet Parasitol*, **144**, 1-9.
- Feng, Y. and Xiao, L.** (2011). Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clin Microbiol Revs*, **24**, 110-140.
- Fichtenbaum CJ, Zackin R, Feinberg J, Benson C, Griffiths JK, Team ACTGNWCS.** (2000) Rifabutin but not clarithromycin prevents Cryptosporidiosis in persons with advanced HIV infection. *AIDS*, **14**(18), 2889-2893.
- Franzén O, Jerlström-Hultqvist J, Castro E, Sherwood E, Ankarklev J, Reiner DS, et al.** (2009). Draft genome sequencing of *Giardia intestinalis* assemblage B isolate GS: is human giardiasis caused by two different species? *PLoS Pathog.* **6**, 1-14.
- Gardner TB, Hill DR.** (2001). Treatment of giardiasis. *Clin Microbiol Rev*, **14**, 114.
- Garner, C. E., Smith, S., Bardhan, P. K., Ratcliffe, N. M. and Probert, C. S.** (2009). A pilot study of faecal volatile organic compounds in faeces from cholera patients in Bangladesh to determine their utility in disease diagnosis. *Trans R Soc Trop Med Hyg*, **103**, 1171-1173.
- Garner, C. E., Smith, S., De Lacy Costello, B., White, P., Spencer, R., Probert, C. S. J. and Ratcliffe, N. M.** (2007). Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *FASEB J*, **21**, 1675-1688.
- Gascón J, Abós R, Valls ME, Corachán M.** (1990). Mebendazole and metronidazole in giardial infections. *Trans R Soc Trop Med Hyg*, **84**, 694.
- Gassama, A., Sow, P. S., Fall, F., Camara, P., Gueye-N'diaye, A., Seng, R., Samb, B., M'Boup, S. and Aidara-Kane, A.** (2001). Ordinary and opportunistic enteropathogens associated with diarrhea in Senegalese adults in relation to human immunodeficiency virus serostatus. *Int J Infect Dis*, **5**, 192-198.

- Gatei W, Ashford RW, Beeching NJ, Kamwati SK, Greensill J, Hart CA.** (2002). *Cryptosporidium muris* infection in an HIV-infected adult, Kenya. *Emerg Infect Dis*, **8**, 204–206.
- Gatei, W., Barrett, D., Lindo, J. F., Eldemire-Shearer, D., Cama, V. and Xiao, L.** (2008). Unique *Cryptosporidium* population in HIV-infected persons, Jamaica. *Emerg Infect Dis*, **14**, 841–843.
- Gatei, W., Das, P., Dutta, P., Sen, A., Cama, V., Lal, A. A. and Xiao, L.** (2007). Multilocus sequence typing and genetic structure of *Cryptosporidium hominis* from children in Kolkata, India. *Infect Genet Evol*, **7**, 197–205.
- Gatei, W., Greensill, J., Ashford, R. W., Cuevas, L. E., Parry, C. M., Cunliffe, N. A., Beeching, N. J. and Hart, C. A.** (2003). Molecular analysis of the 18S rRNA gene of *Cryptosporidium* parasites from patients with or without human immunodeficiency virus infections living in Kenya, Malawi, Brazil, the United Kingdom, and Vietnam. *J Clin Microbiol*, **41**, 1458–1462.
- Gatei, W., Hart, C. A., Gilman, R. H., Das, P., Cama, V. and Xiao, L.** (2006). Development of a multilocus sequence typing tool for *Cryptosporidium hominis*. *J Eukaryot Microbiol*, **53** Suppl 1, S43–48.
- Gatei, W., Suputtamongkol, Y., Waywa, D., Ashford, R. W., Bailey, J. W., Greensill, J., Beeching, N. J. and Hart, C. A.** (2002). Zoonotic species of *Cryptosporidium* are as prevalent as the anthroponotic in HIV-infected patients in Thailand. *Ann Trop Med Parasitol*, **96**, 797–802.
- Gatei W, Wamae CN ,Mbae C, Waruru A, Mulinge E, Waithera T,et al.** (2006). Cryptosporidiosis :prevalence, genotype analysis, and symptoms associated with infections in children in Kenya. *Am J Trop Med Hyg*, **75**, 78–82.
- GBD 2015.** (2016). Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*, **388**, 1459–1544.

- Geurden, T., Geldhof, P., Levecke, B., Martens, C., Berkvens, D., Casaert, S., Vercruysse, J. and Claerebout, E.** (2008). Mixed *Giardia duodenalis* assemblage A and E infections in calves. *Int J Parasitol*, **38**, 259-264.
- Geurden, T., Goossens, E., Levecke, B., Vercammen, F., Vercruysse, J. and Claerebout, E.** (2009a). Occurrence and molecular characterization of *Cryptosporidium* and *Giardia* in captive wild ruminants in Belgium. *J Zoo Wildl Med*, **40**, 126-130.
- Geurden, T., Levecke, B., Cacció, S. M., Visser, A., de Groote, G., Casaert, S., Vercruysse, J. and Claerebout, E.** (2009b). Multilocus genotyping of *Cryptosporidium* and *Giardia* in non-outbreak related cases of diarrhoea in human patients in Belgium. *Parasitol*, **136**, 1161-1168.
- Ghaffari, S. and Kalantari, N.** (2014). A multi-locus study of *Cryptosporidium* parasites isolated from patients living in Iran, Malawi, Nigeria, the United Kingdom, and Vietnam. *Iran J Parasitol*, **9**, 79-89.
- Giles, M., Webster, K. A., Marshall, J. A., Catchpole, J. and Goddard, T. M.** (2001). Experimental infection of a lamb with *Cryptosporidium parvum* genotype 1. *Vet Rec*, **149**, 523-525.
- Gillis, J. C.; Wiseman, L. R.** (1996). Secnidazole. A review of its antimicrobial activity, pharmacokinetic properties and therapeutic use in the management of protozoal infections and bacterial vaginosis. *Drugs*, **51** (4): 621-38.
- Glaberman, J.E. Moore, C.J. Lowery, R.M. Chalmers, I. Sulaiman, K. Elwin, P.J. Rooney, B.C. Millar, J.S. Dooley, A.A. Lal and L. Xiao.** (2002). Three drinking-water-associated cryptosporidiosis outbreaks, Northern Ireland. *Emerg. Infect. Dis*, **8**, 631-633.
- Goldin AJ, Hall A, Sarker RN, et al.** (1993). Diagnosis of *Giardia duodenalis* infection in Bangladeshi infants: faecal antigen capture ELISA. *Trans R Soc Trop Med Hyg*, **87**, 428.
- Graczyk, T. K. and Fried, B.** (2007). Human waterborne trematode and protozoan infections. *Adv Parasitol*, **64**, 111-160.

- Greene, C. E., Jacobs, G. J. and Prickett, D.** (1990). Intestinal malabsorption and cryptosporidiosis in an adult dog. *J Am Vet Med Assoc*, **197**, 365-367.
- Greenwood, David.** (2008). Antimicrobial drugs : chronicle of a twentieth century medical triumph (1. publ. ed.). *Oxford: Oxford University Press*, 239.
- Gross TL, Wheat J, Bartlett M, O'Connor KW.** (1986). AIDS and multiple system involvement with *Cryptosporidium*. *Am J Gastroenterol*, **81**, 456.
- Gyorkos, T. W., MacLean, J. D., Viens, P., Chheang, C. and Kokoskin-Nelson, E.** (1992). Intestinal parasite infection in the Kampuchean refugee population 6 years after resettlement in Canada. *J Infect Dis*, **166**, 413-417.
- Haghighi, A., Khorashad, A. S., Nazemalhosseini Mojarad, E., Kazemi, B., Rostami Nejad, M. and Rasti, S.** (2009). Frequency of enteric protozoan parasites among patients with gastrointestinal complaints in medical centers of Zahedan, Iran. *Trans R Soc Trop Med Hyg*, **103**, 452-454.
- Hamilton, Richart** (2015). Tarascon Pocket Pharmacopoeia 2015 Deluxe Lab-Coat Edition. Jones & Bartlett Learning, 50. ISBN 9781284057560.
- Hanevik K, Hausken T, Morken MH, et al.** (2007). Persisting symptoms and duodenal inflammation related to *Giardia duodenalis* infection. *J Infect*, **55**, 524.
- Hanevik K, Mørch K, Eide GE, et al.** (2008). Effects of albendazole/metronidazole or tetracycline/folate treatments on persisting symptoms after *Giardia* infection: a randomized open clinical trial. *Scand J Infect Dis*, **40**, 517.
- Harp, J. A.** (2003). *Cryptosporidium* and host resistance: historical perspective and some novel approaches. *Anim Health Res Rev*, **4**, 53-62.
- Harter, L., Frost, F., Grunenfelder, G., Perkins-Jones, K. and Libby, J.** (1984). Giardiasis in an Infant and Toddler Swim Class. *Am J Public Health*, **74**, 155-156.

- Hill, D. R.** (2001). *Giardia lamblia*. In *Principles and practice of clinical parasitology* (eds. Gillespie, S. H., and Pearson, R. D.), John Wiley & Sons Ltd, Chichester; UK.
- Hira, K. G., Mackay, M. R., Hempstead, A. D., Ahmed, S., Karim, M. M., O'Connor, R. M., Hibberd, P. L., Calderwood, S. B., Ryan, E. T., Khan, W. A. and Ward, H. D.** (2011). Genetic diversity of *Cryptosporidium* spp. from Bangladeshi children. *J Clin Microbiol*, **49**, 2307-2310.
- Holmberg SD, Moorman AC, Von Bargen JC, et al.** (1998). Possible effectiveness of clarithromycin and rifabutin for cryptosporidiosis chemoprophylaxis in HIV disease. HIV Outpatient Study (HOPS). Investigators. *JAMA*, **279**(5), 384-386.
- Homan, W. L. and Mank, T. G.** (2001). Human giardiasis: genotype linked differences in clinical symptomatology. *Int J Parasitol*, **31**, 822-826.
- Hopkins, R. M., Meloni, B. P., Groth, D. M., Wetherall, J. D., Reynoldson, J. A. and Thompson, R. C. A.** (1997). Ribosomal RNA sequencing reveals differences between the genotypes of *Giardia* isolates recovered from humans and dogs living in the same locality. *J Parasitol*, **83**, 44-51.
- Hoque ME, Hope VT, Scragg R, Kjellström T, Lay-Yee R.** (2001). Nappy handling and risk of giardiasis. *Lancet*, 357 (9261), 1017-1018.
- <http://www.dpd.cdc.gov/dpdx/cryptosporidiosis>.
- <https://en.wikipedia.org/wiki/Metronidazole>
- <http://www.nhs.uk/Conditions/Giardiasis/Pages/Prevention.aspx>
- <http://www.health.state.mn.us/divs/idepc/diseases/cryptosporidiosis/prevention.html><https://aidsinfo.nih.gov/guidelines/html/4/adult-and-adolescent-oi-prevention-and-treatment-guidelines/323/cryptosporidia>

https://www.researchgate.net/post/What_are_the_differences_between_germplasm_accession_genotype_and_population_and_also_between_heirloom_and_landrace [accessed Aug 5, 2017].

<http://www.k-state.edu/parasitology/basicbio>

<https://data.unicef.org/topic/child-health/diarrhoeal-disease/#>

<https://www.uptodate.com/contents/epidemiology-clinical-manifestations-and-diagnosis-of-cryptosporidiosis...> Leder, 2016 Epidemiology, clinical manifestations, and diagnosis of cryptosporidiosis

Huang, D. B. and White, A. C. (2006). An updated review on *Cryptosporidium* and *Giardia*. *Gastroenterol Clin North Am*, **35**, 291-314.

Huey CS, Mahdy MA, Al-Mekhlafi HM, Nasr NA, Lim YA, Mahmud R, Surin J. (2013). Multilocus genotyping of *Giardia duodenalis* in Malaysia. *Infect Genet Evol*, **17**, 269-276

Hunter, P. R. and Thompson, R. C. (2005). The zoonotic transmission of *Giardia* and *Cryptosporidium*. *Int J Parasitol*, **35**, 1181-1190.

Hussein AI, Yamaguchi T, Nakamoto K, Iseki M, Tokoro M. (2009). Multiple subgenotype infections of *Giardia intestinalis* detected in Palestinian clinical cases using a subcloning approach. *Parasitol Int*, **58**, 258-62.

H. V. Huynh; J. H. H. Ho; T. C. Neo; L. L. Koh. (2005). Solvent-controlled selective synthesis of a trans-configured benzimidazoline-2-ylidene palladium(II) complex and investigations of its Heck-type catalytic activity. *J Organomet Chem*, **690** (16), 3854-3860.

Ignatius, R., Gahutu, J. B., Klotz, C., Steininger, C., Shyirambere, C., Lyng, M., Musemakweri, A., Aebischer, T., Martus, P., Harms, G. and Mockenhaupt, F. P. (2012). High prevalence of *Giardia duodenalis* Assemblage B infection and association with underweight in Rwandan children. *PLoS Negl Trop Dis*, **6**, e1677.

International Drug Price Indicator Guide. Rifabutin. Retrieved 8 December 2016.

Iqbal, J., Hira, P. R., Al-Ali, F. and Philip, R. (2001). Cryptosporidiosis in Kuwaiti children: seasonality and endemicity. *Clin Microbiol Infect*, **7**, 261-266.

Janoff, E. N., Mead, P. S., Mead, J. R., Echeverria, P., Bodhidatta, L., Bhaibulaya, M., Sterling, C. R. and Taylor, D. N. (1990). Endemic *Cryptosporidium* and *Giardia lamblia* infections in a Thai orphanage. *Am J Trop Med Hyg*, **43**, 248-256.

Javier Enriquez, F., Avila, C. R., Ignacio Santos, J., Tanaka-Kido, J., Vallejo, O. and Sterling, C. R. (1997). *Cryptosporidium* infections in Mexican children: clinical, nutritional, enteropathogenic, and diagnostic evaluations. *Am J Trop Med Hyg*, **56**, 254-257.

Jeon, H. K., Yong, T. S., Sohn, W. M., Chai, J. Y., Hong, S. J., Han, E. T., Jeong, H. G., Chhakda, T., Sinuon, M., Socheat, D. and Eom, K. S. (2011). Molecular identification of *Taenia* tapeworms by Cox1 gene in Koh Kong, Cambodia. *Korean J Parasitol*, **49**, 195-197.

Jex, A. R. and Gasser, R. B. (2008). Analysis of the genetic diversity within *Cryptosporidium hominis* and *Cryptosporidium parvum* from imported and autochthonous cases of human cryptosporidiosis by mutation scanning. *Electrophoresis*, **29**, 4119-4129.

Johal S. S., Hammond J., Wight N. (2002). Role of flexible sigmoidoscopy in *C. difficile*-associated diarrhoea (CDAD). *Gut*, **51**(suppl 3), 414.

Jongwutiwes, S., Tiangtip, R., Yentakarm, S. and Chantachum, N. (2002). Simple method for long-term copro-preservation of *Cryptosporidium* oocysts for morphometric and molecular analysis. *Trop Med Int Health*, **7**, 257-264.

Johnston AR, Gillespie TR, Rwego IB, McLachlan TL, Kent AD, Goldberg TL. (2010). Molecular epidemiology of cross-species *Giardia duodenalis* transmission in western Uganda. *PLoS Negl Trop Dis*. **4**, e683.

Jokipii L, Jokipii AM. (1986). Timing of symptoms and oocyst excretion in human cryptosporidiosis. *N Engl J Med*, **315**, 1643.

- Jongwutiwes, S., Tiangtip, R., Yentakarm, S. and Chantachum, N.** (2002). Simple method for long-term copro-preservation of *Cryptosporidium* oocysts for morphometric and molecular analysis. *Trop Med Int Health*, **7**, 257-264.
- Juranek DD.** Cryptosporidiosis. In: Strickland GT, editor. Hunter's Tropical Medicine and Emerging Infectious Diseases,. 8th ed. Philadelphia: WB Saunders; 2000.
- Kappus, K. D., Lundgren Jr, R. G., Juranek, D. D., Roberts, J. M. and Spencer, H. C.** (2004). Intestinal parasitism in the United States: Update on a continuing problem. *Am J Trop Med Hyg*, **50**, 705-713.
- Karanis, P., Kourenti, C. and Smith, H.** (2007). Waterborne transmission of protozoan parasites: A worldwide review of outbreaks and lessons learnt. *J Water Health*, **5**, 1-38.
- Katellaris PH, Farthing MJG.** (1992). Diarrhea and malabsorption in giardiasis: a multifactorial process? *Gut*, **33**, 295-297.
- Katsumata, T., Hosea, D., Ranuh, I. G., Uga, S., Yanagi, T. and Kohno, S.** (2000). Short report: possible *Cryptosporidium muris* infection in humans. *Am J Trop Med Hyg*, **62**, 70-72.
- Katzung, B. G.** (2004). Basic and Clinical Pharmacology (9th ed.).
- Khieu, V., Schar, F., Forrer, A., Hattendorf, J., Marti, H., Duong, S., Vounatsou, P., Muth, S. and Odermatt, P.** (2014a). High prevalence and spatial distribution of *Strongyloides stercoralis* in rural Cambodia. *PLoS Negl Trop Dis*, **8**, e2854.
- Khieu, V., Schar, F., Marti, H., Bless, P. J., Char, M. C., Muth, S. and Odermatt, P.** (2014b). Prevalence and risk factors of *Strongyloides stercoralis* in Takeo Province, Cambodia. *Parasit Vectors*, **7**, 221.
- Kjos, S. A., Jenkins, M., Okhuysen, P. C. and Chappell, C. L.** (2005). Evaluation of recombinant oocyst protein CP41 for detection of *Cryptosporidium*-specific antibodies. *Clin Diagn Lab Immunol*, **12**, 268-272.

- Koch, K. L., Phillips, D. J., Aber, R. C. and Current, W. L.** (1985). Cryptosporidiosis in hospital personnel. Evidence for person-to-person transmission. *Ann Intern Med*, **102**, 593-596.
- Kohli A, Bushen OY, Pinkerton RC, Houpt E, Newman RD, Sears CL, et al.** (2008). *Giardia duodenalis* assemblage, clinical presentation and markers of intestinal inflammation in Brazilian children. *Trans R Soc Trop Med Hyg*, **102**, 718-25.
- Korich DG, Mead JR, Madore MS, Sinclair NA, Sterling CR.** (1990). Effects of ozone, chlorine dioxide, chlorine, and monochloramine on *Cryptosporidium parvum* oocyst viability. *Appl Environ Microbiol*, **56** (5), 1423-8.
- Kosek, M., Bern, C. and Guerrant, R. L.** (2003). The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull World Health Organ*, **81**, 197-204.
- Kosuwin R, Putaporntip C, Pattanawong U, Jongwutiwes S.** (2010). Clonal diversity in *Giardia duodenalis* isolates from Thailand: evidences for intragenic recombination and purifying selection at the beta giardin locus. *Gene*, **449**, 1-8.44.
- Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al.** (2013). Burden and etiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective
- Krause PJ, Lepore T, Sikand VK, et al.** (2000). Atovaquone and azithromycin for the treatment of babesiosis. *N. Engl. J. Med.* **343** (20), 1454-8.
- Krumkamp R, Sarpong N, Schwarz NG, Adelkofer J, Loag W, Eibach D, et al.** (2015). Gastrointestinal Infections and Diarrheal Disease in Ghanaian Infants and Children: An Outpatient Case-Control Study. *PLoS Neglect Trop D*, **9**, e0003568.
- Lalle, M., Pozio, E., Capelli, G., Bruschi, F., Crotti, D. and Cacciò, S. M.** (2005). Genetic heterogeneity at the β -giardin locus among human and animal isolates of *Giardia duodenalis* and identification of potentially zoonotic subgenotypes. *Int J Parasitol*, **35**, 207-213.

- Lalle M, Bruschi F, Castagna B, Campa M, Pozio E, Caccio SM.** (2009). High genetic polymorphism among *Giardia duodenalis* isolates from Sahrawi children. *Trans R Soc Trop Med Hyg*, **103**, 834-8.
- Leav, B. A., Mackay, M. R., Anyanwu, A., RM, O. C., Cevallos, A. M., Kindra, G., Rollins, N. C., Bennish, M. L., Nelson, R. G. and Ward, H. D.** (2002). Analysis of sequence diversity at the highly polymorphic Cpgp40/15 locus among *Cryptosporidium* isolates from human immunodeficiency virus-infected children in South Africa. *Infect Immun*, **70**, 3881-3890.
- Lebbad M, Ankarklev J, Tellez A, Leiva B, Andersson JO, Svard S.** (2008). Dominance of *Giardia* assemblage B in Leon, Nicaragua. *Acta Trop*, 106, 44-53.
- Lebbad M, Petersson I, Karlsson L, Botero-Kleiven S, Andersson JO, Svenungsson B, et al.** (2011). Multilocus genotyping of human *Giardia* isolates suggests limited zoonotic transmission and association between assemblage B and flatulence in children. *PLoS Negl Trop Dis*, **5**, e1262.
- Lebbad, M., Petersson, I., Karlsson, L., Botero-Kleiven, S., Andersson, J. O., Svenungsson, B. and Svärd, S. G.** (2011). Multilocus genotyping of human *Giardia* isolates suggests limited zoonotic transmission and association between assemblage B and flatulence in children. *PLoS Negl Trop Dis*, **5**.
- Lee, J. H., Kang, M., Choung, S. J., Ogino, K., Miyata, S., Kim, M. S., Park, J. Y. and Kim, J. B.** (2004). The preparation of TiO₂ nanometer photocatalyst film by a hydrothermal method and its sterilization performance for *Giardia lamblia*. *Water Res*, **38**, 713-719.
- Lee, K. J., Bae, Y. T., Kim, D. H., Deung, Y. K., Ryang, Y. S., Kim, H. J., Im, K. I. and Yong, T. S.** (2002). Status of intestinal parasites infection among primary school children in Kampongcham, Cambodia. *Korean J Parasitol*, **40**, 153-155.
- Lemke, Thomas L.; Williams, David A.** (2008). Foye's Principles of Medicinal Chemistry. Lippincott Williams & Wilkins. p. 675.

- Leoni, F., Amar, C., Nichols, G., Pedraza-Diaz, S. and McLauchlin, J.** (2006a). Genetic analysis of *Cryptosporidium* from 2414 humans with diarrhoea in England between 1985 and 2000. *J Med Microbiol*, **55**, 703-707.
- Leoni, F., Gallimore, C.I., Green, J. and McLauchlin, J.** (2006b). Characterisation of small double stranded RNA molecule in *Cryptosporidium hominis*, *Cryptosporidium felis* and *Cryptosporidium meleagridis*. *Parasitol Int*, **55**, 299-306.
- Leoni, F., Mallon, M.E., Smith, H.V., Tait, A. and McLauchlin, J.** (2007). Multilocus analysis of *Cryptosporidium hominis* and *Cryptosporidium parvum* isolates from sporadic and outbreak-related human cases and *C. parvum* isolates from sporadic livestock cases in the United Kingdom. *J Clin Microbiol*, **45**, 3286-3294.
- Levine M M, Robins-rowne RM.** (2012). Factors that explain excretion of enteric pathogens by persons without diarrhea. *Clin Infect Dis*, **55** (Suppl4), S303-311.
- Li, W. H.** (1997). Molecular Evolution. Sinauer, Sunderland, MA
- Lim YA, Iqbal A, Surin J, Sim BL, Jex AR, Nolan MJ, Smith HV, Gasser RB.** (2011). First genetic classification of *Cryptosporidium* and *Giardia* from HIV/AIDS patients in Malaysia. *Infect Genet Evol.* **11**, 968-974.
- Lima, A. A. and Guerrant, R. L.** (2007). Heavy cryptosporidial infections in children in northeast Brazil: comparison of *Cryptosporidium hominis* and *Cryptosporidium parvum*. *Trans R Soc Trop Med Hyg*, **101**, 378-384.
- Liu, H., Shen, Y., Yin, J., Yuan, Z., Jiang, Y., Xu, Y., Pan, W., Hu, Y. and Cao, J.** (2014). Prevalence and genetic characterization of *Cryptosporidium*, *Enterocytozoon*, *Giardia* and *Cyclospora* in diarrheal outpatients in China. *BMC Infect Dis*, **14**, 25.
- Liu, J, Platts-Mills, JA, Juma, J et al.** (2016). Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control. *Lancet*, **388**, 1291-1301.

- Lobo ML, Augusto J, Antunes F, Ceita J, Xiao L, Codices V, et al.** (2014). *Cryptosporidium spp., Giardia duodenalis, Enterocytozoon bieneusi* and other intestinal parasites in young children in Lobata province, Democratic Republic of Sao Tome and Principe. *PLoS One*, **9**(5), e97708
- Lurio, J., Verson, H. and Karp, S.** (1991). Intestinal parasites in Cambodians: comparison of diagnostic methods used in screening refugees with implications for treatment of populations with high rates of infestation. *J Am Board Fam Pract*, **4**, 71-78.
- Machado RS, Silva MR, Viriato A.** (2008). Furazolidone, tetracycline and omeprazole: a low-cost alternative for *Helicobacter pylori* eradication in children. *J Pediatr (Rio J)*, **84**(2), 160-5.
- Mahmud MA, Chappell CL, Hossain MM, Huang DB, Habib M, DuPont HL.** (2001). Impact of breast feeding on *Giardia lamblia* infections in Bilbeis, Egypt. *Am J Trop Med Hyg*, **65**, 257-260.
- Mak, J. W.** (2004). Important zoonotic intestinal protozoan parasites in Asia. *Trop Biomed*, **21**, 39-50.
- Mallon, M. E., MacLeod, A., Wastling, J. M., Smith, H. and Tait, A.** (2003). Multilocus genotyping of *Cryptosporidium parvum* Type 2: population genetics and sub-structuring. *Infection, Genetics And Evolution*, **3**, 207-218.
- Mank, T. G., Zaat, J. O. M., Deelder, A. M., Van Eijk, J. T. M. and Polderman, A. M.** (1997). Sensitivity of microscopy versus enzyme immunoassay in the laboratory diagnosis of giardiasis. *European Journal of Clinical Microbiology and Infectious Diseases*, **16**, 615-619.
- Manser, M., Granlund, M., Edwards, H., Saez, A., Petersen, E., Evengard, B., Chiodini, P., European Society of Clinical, M. and Infectious Diseases Study Group on Clinical, P.** (2014). Detection of *Cryptosporidium* and *Giardia* in clinical laboratories in Europe--a comparative study. *Clin Microbiol Infect*, **20**, 65-71.
- Marmot, M.** (2008). Social determinants of health, Health in Asia and the Pacific.
- Marshall A. W, DeSouza M, Morgan M. Y.** (1985). Plasma 3-methylbutanal in man and its relationship to hepatic encephalopathy. *Clin Physiol*, **5**, 53-62.

- Marshall, M. M., Naumovitz, D., Ortega, Y. and Sterling, C. R.** (1997). Waterborne protozoan pathogens. *Clin Microbiol Rev*, **10**, 67-85.
- Mbae, C. K., Nokes, D. J., Mulinge, E., Nyambura, J., Waruru, A. and Kariuki, S.** (2013). Intestinal parasitic infections in children presenting with diarrhoea in outpatient and inpatient settings in an informal settlement of Nairobi, Kenya. *BMC Infect Dis*, **13**, 243.
- McElligott, J. T., Naaktgeboren, C., Makuma-Massa, H., Summer, A. P. and Deal, J. L.** (2013). Prevalence of intestinal protozoa in communities along the Lake Victoria region of Uganda. *Int J Infect Dis*, **17**, e658-659.
- Meamar, A. R., Guyot, K., Certad, G., Dei-Cas, E., Mohraz, M., Mohebbi, M., Mohammad, K., Mehdad, A. A., Rezaie, S. and Rezaian, M.** (2007). Molecular characterization of *Cryptosporidium* isolates from humans and animals in Iran. *Appl Environ Microbiol*, **73**, 1033-1035.
- Meloni, B. P., Lymbery, A. J. and Thompson, R. C.** (1995). Genetic characterization of isolates of *Giardia duodenalis* by enzyme electrophoresis: implications for reproductive biology, population structure, taxonomy, and epidemiology. *J Parasitol*, **81**, 368-383.
- Mills JA, Gratz J, Mduma E, Svensen E, Amour C, et al.** (2014). Association between stool enteropathogen quantity and disease in Tanzanian children using Taq Manarray cards: a nested case-control study. *Am J Trop Med Hyg*, **90**, 133-138.
- Minetti, C., Taweean, W., Hogg, R., Featherstone, C., Randle, N., Latham, S. M. and Wastling, J. M.** (2014). Occurrence and diversity of *Giardia duodenalis* assemblages in livestock in the UK. *Transbound Emerg Dis*, **61**, e60-67.
- Miyamoto, K., Kirinoki, M., Matsuda, H., Hayashi, N., Chigusa, Y., Sinuon, M., Chuor, C. M. and Kitikoon, V.** (2014). Field survey focused on *Opisthorchis viverrini* infection in five provinces of Cambodia. *Parasitol Int*, **63**, 366-373.

- Mohammed MAK, Lim YA, Surin J, Wan KL, Al-Mekhlafi MS.** (2008). Risk factors for endemic giardiasis: highlighting the possible association of contaminated water and food. *Trans R Soc Trop Med Hyg*, **102** (5), 465-470.
- Molbak K, Hojlyng N, A, Sa JC, Ingholt L, et al.** (1993). Cryptosporidiosis in infancy and childhood mortality in Guinea Bissau, west Africa. *Bmj*, **307**, 417-420
- Molloy SF, Smith HV, Kirwan P, Nichols RA, Asaolu SO, Connelly L, et al.** (2010). Identification of a high diversity of *Cryptosporidium* species genotypes and subtypes in a pediatric population in Nigeria. *Am J Trop Med Hyg*, **82**(4), 608-13.
- olloy SF, Tanner CJ, Kirwan P, Asaolu SO, Smith HV, Nichols RA, et al.** (2011). Sporadic *Cryptosporidium* infection in Nigerian children: risk factors with species identification. *Epidemiol Infect*, **139**, 946-954
- Monis, P. T., Andrews, R. H., Mayrhofer, G. and Ey, P. L.** (2003). Genetic diversity within the morphological species *Giardia intestinalis* and its relationship to host origin. *Infect Genet Evol*, **3**, 29-38.
- Moore Catrin E, Kristin Elwin, Nget Phot , Chanthou Seng, Saroeun Mao, Kuong Suy, Varun Kumar, Johanna Nade Rachel Bousfield, Sanuki Perera, Wendi Bailey, Nicholas J. Beeching, Christopher M. Parry, Rachel M. Chalmers,** (2016). Molecular Characterization of *Cryptosporidium* Species and *Giardia duodenalis* from Symptomatic Cambodian Children. *PLoS Negl Trop Dis*, **7**, 10(7).
- Moore, C. E., Hor, P. C., Soeng, S., Sun, S., Lee, S. J., Parry, C. M., Day, N. P. and Stoesser, N.** (2012). Changing patterns of gastrointestinal parasite infections in Cambodian children: 2006-2011. *J Trop Pediatr*, **58**, 509-512.
- Moore, C. E., Nget, P., Saroeun, M., Kuong, S., Chanthou, S., Kumar, V., Bousfield, R., Nader, J., Bailey, J. W., Beeching, N. J., Day, N. P. and Parry, C. M.** (2015). Intestinal parasite infections in symptomatic children attending hospital in siem reap, cambodia. *PLoS One*, **10**, 5.

- Mørch K, Hanevik K, Robertson LJ, et al.** (2008). Treatment-ladder and genetic characterisation of parasites in refractory giardiasis after an outbreak in Norway. *J Infect*, **56**, 268.
- Morgan, U., Weber, R., Xiao, L., Sulaiman, I., Thompson, R. C., Ndiritu, W., Lal, A., Moore, A. and Deplazes, P.** (2000a). Molecular characterization of *Cryptosporidium* isolates obtained from human immunodeficiency virus-infected individuals living in Switzerland, Kenya, and the United States. *J Clin Microbiol*, **38**, 1180-1183.
- Morgan, U., Xiao, L., Sulaiman, I., Weber, R., Lal, A. A., Thompson, R. C. and Deplazes, P.** (1999). Which genotypes/species of *Cryptosporidium* are humans susceptible to? *J Eukaryot Microbiol*, **46**, 42S-43S.
- Morgan, U. M., Xiao, L., Fayer, R., Lal, A. A. and Thompson, R. C.** (2000b). Epidemiology and strain variation of *Cryptosporidium parvum*. *Contrib Microbiol*, **6**, 116-139.
- Morgan, U. M., Xiao, L., Hill, B. D., O'Donoghue, P., Limor, J., Lal, A. and Thompson, R. C.** (2000c). Detection of the *Cryptosporidium parvum* "human" genotype in a dugong (*Dugong dugon*). *J Parasitol*, **86**, 1352-1354.
- Morrison, L. J., Mallon, M. E., Smith, H. V., MacLeod, A., Xiao, L. and Tait, A.** (2008). The population structure of the *Cryptosporidium parvum* population in Scotland: a complex picture. *Infect Genet Evol*, **8**, 121-129.
- Mor SM, Tzipori S** (2008). Cryptosporidiosis in children in Sub-Saharan Africa: a lingering challenge. *Clin Infect Dis*, **47**, 915-921
- Morse, T. D., Nichols, R. A., Grimason, A. M., Campbell, B. M., Tembo, K. C. and Smith, H. V.** (2007). Incidence of cryptosporidiosis species in paediatric patients in Malawi. *Epidemiol Infect*, **135**, 1307-1315.
- Moyo, S. J., Gro, N., Matee, M. I., Kitundu, J., Myrmel, H., Mylvaganam, H., Maselle, S. Y. and Langeland, N.** (2011). Age specific aetiological agents of diarrhoea in hospitalized children aged less than five years in Dar es Salaam, Tanzania. *BMC Pediatr*, **11**, 19.

- Mukherjee AK, Chowdhury P, Bhattacharya MK, Ghosh M, Rajendran K, Ganguly S.** (2009). Hospital-based surveillance of enteric parasites in Kolkata. *BMC Res Notes*, **2**, 110.
- Mukhopadhyay, C., Wilson, G., Pradhan, D. and Shivananda, P. G.** (2007). Intestinal protozoal infestation profile in persistent diarrhea in children below age 5 years in western Nepal. *Southeast Asian J Trop Med Public Health*, **38**, 13-19.
- Mungthin, M., Suwannasaeng, R., Naaglor, T., Areekul, W. and Leelayoova, S.** (2001). Asymptomatic intestinal microsporidiosis in Thai orphans and child-care workers. *Trans R Soc Trop Med Hyg*, **95**, 304-306.
- Murray, Patrick R., Ken S. Rosenthal, and Michael A. Pfaller.** Medical Microbiology. 5th ed. Philadelphia: Elsevier Inc., 2005: 855-856.
- Musher DM, Logan N, Bressler AM, et. al.** (2009). Nitazoxanide versus vancomycin in *Clostridium difficile* infection: a randomized, double-blind study. *Clin Infect Dis*, **48**(4), e41-6.
- Muthusamy, D., Rao, S. S., Ramani, S., Monica, B., Banerjee, I., Abraham, O. C., Mathai, D. C., Primrose, B., Muliyl, J., Wanke, C. A., Ward, H. D. and Kang, G.** (2006). Multilocus genotyping of *Cryptosporidium* sp. isolates from human immunodeficiency virus-infected individuals in South India. *J Clin Microbiol*, **44**, 632-634.
- Mwachari, C., Batchelor, B. I., Paul, J., Waiyaki, P. G. and Gilks, C. F.** (1998). Chronic diarrhoea among HIV-infected adult patients in Nairobi, Kenya. *J Infect*, **37**, 48-53.
- Nain CK, Dutt P, Vinayak VK.** 1991. Alterations in enzymatic activities of the intestinal mucosa during the course of *Giardia lamblia* infection in mice. *Ann Trop Med Parasitol*, **85**, 515-522.
- National Statistical Office (Malawi) Zomba, Malawi.** (2001). Malawi Demographic and Health Survey, 2000.

- Newman, R. D., Sears, C. L., Moore, S. R., Nataro, J. P., Wuhib, T., Agnew, D. A., Guerrant, R. L. and Lima, A. A.** (1999). Longitudinal study of *Cryptosporidium* infection in children in northeastern Brazil. *J Infect Dis*, **180**, 167-175.
- Nichols, R. A., Moore, J. E. and Smith, H. V.** (2006). A rapid method for extracting oocyst DNA from *Cryptosporidium*-positive human faeces for outbreak investigations. *J Microbiol Methods*, **65**, 512-524.
- Nkenfou, C. N., Nana, C. T. and Payne, V. K.** (2013). Intestinal parasitic infections in HIV infected and non-infected patients in a low HIV prevalence region, West-Cameroon. *PLoS One*, **8**, e57914.
- Nuchjangreed C1, Boonrod K, Ongerth J, Karanis P.** (2008). Prevalence and molecular characterization of human and bovine *Cryptosporidium* isolates in Thailand. *Parasitol Res*, **103**(6), 1347-53.
- O'Connor, R. M., Wanyiri, J. W., Cevallos, A. M., Priest, J. W. and Ward, H. D.** (2007). *Cryptosporidium parvum* glycoprotein gp40 localizes to the sporozoite surface by association with gp15. *Mol Biochem Parasitol*, **156**, 80-83.
- O'Handley, R. M., Buret, A. G., McAllister, T. A., Jelinski, M. and Olson, M. E.** (2001). Giardiasis in dairy calves: effects of fenbendazole treatment on intestinal structure and function. *Int J Parasitol*, **31**, 73-79.
- Odiit, A., Mulindwa, A., Nalumansi, E., Mphahlele, M. J., Seheri, L. M., Mwenda, J. M. and Kisakye, A.** (2014). Rotavirus prevalence and genotypes among children younger than 5 years with acute diarrhea at Mulago National Referral Hospital, Kampala, Uganda. *Pediatr Infect Dis J*, **33 Suppl 1**, S41-44.
- Okhuysen, P. C. and Chappell, C. L.** (2002). *Cryptosporidium* virulence determinants-are we there yet? *Int J Parasitol*, **32**, 517-525.

- Okhuysen, P. C., Chappell, C. L., Crabb, J. H., Sterling, C. R. and DuPont, H. L.** (1999). Virulence of three distinct *Cryptosporidium parvum* isolates for healthy adults. *J Infect Dis*, **180**, 1275-1281.
- Ong, C. S., Eisler, D. L., Alikhani, A., Fung, V. W., Tomblin, J., Bowie, W. R. and Isaac-Renton, J. L.** (2002). Novel *cryptosporidium* genotypes in sporadic cryptosporidiosis cases: first report of human infections with a cervine genotype. *Emerg Infect Dis*, **8**, 263-268.
- Palmer, C. J., Xiao, L., Terashima, A., Guerra, H., Gotuzzo, E., Saldias, G., Bonilla, J. A., Zhou, L., Lindquist, A. and Upton, S. J.** (2003). *Cryptosporidium muris*, a rodent pathogen, recovered from a human in Peru. *Emerg Infect Dis*, **9**, 1174-1176.
- Pavone, R., Schinaia, N., Hart, C. A., Getty, B., Molyneux, M. and Borgstein, A.** (1990). Viral gastro-enteritis in children in Malawi. *Ann Trop Paediatr*, **10**, 15-20.
- Pawliszyn J.** Theory of solid-phase microextraction. *J Chromatogr Sci* 2000; 38: 270-278.
- Pedraza-Diaz, S., Amar, C. F., McLauchlin, J., Nichols, G. L., Cotton, K. M., Godwin, P., Iversen, A. M., Milne, L., Mulla, J. R., Nye, K., Panigrahl, H., Venn, S. R., Wiggins, R., Williams, M. and Youngs, E. R.** (2001). *Cryptosporidium meleagridis* from humans: molecular analysis and description of affected patients. *J Infect*, **42**, 243-250.
- Peng, M. M., Matos, O., Gatei, W., Das, P., Stantic-Pavlinic, M., Bern, C., Sulaiman, I. M., Glaberman, S., Lal, A. A. and Xiao, L.** (2001). A comparison of *Cryptosporidium* subgenotypes from several geographic regions. *J Eukaryot Microbiol*, **Suppl**, 28S-31S.
- Peng, M. M., Wilson, M. L., Holland, R. E., Meshnick, S. R., Lal, A. A. and Xiao, L.** (2003). Genetic diversity of *Cryptosporidium* spp. in cattle in Michigan: implications for understanding the transmission dynamics. *Parasitol Res*, **90**, 175-180.
- Perez Cordon, G., Cordova Paz Soldan, O., Vargas Vasquez, F., Velasco Soto, J. R., Sempere Bordes, L., Sanchez Moreno, M. and Rosales, M. J.** (2008). Prevalence of enteroparasites and genotyping of *Giardia lamblia* in Peruvian children. *Parasitol Res*, **103**, 459-465.

- Petersen C.** (1995). *Cryptosporidium* and the food supply. *Lancet*, **345**, 1128.
- PharmD, Benjamin Barner, PharmD, Brett Feret.** (2010). Rifaximin: A nonabsorbable, broad-spectrum antibiotic for reduction in the risk for recurrence of overt hepatic encephalopathy.
- Pieniazek, N. J., Bornay-Llinares, F. J., Slemenda, S. B., da Silva, A. J., Moura, I. N., Arrowood, M. J., Ditrich, O. and Addiss, D. G.** (1999). New *cryptosporidium* genotypes in HIV-infected persons. *Emerg Infect Dis*, **5**, 444-449.
- Pirestani, M., Sadraei, J., Dalimi Asl, A., Zavvar, M. and Vaeznia, H.** (2008). Molecular characterization of *Cryptosporidium* isolates from human and bovine using 18s rRNA gene in Shahriar county of Tehran, Iran. *Parasitol Res*, **103**, 467-472.
- Probert, C. S. J., Ahmed, F., Khalid, T., Johnson, E., Smith, S. and Ratcliffe, N.** (2009). Volatile organic compounds as diagnostic biomarkers in gastrointestinal and liver diseases. *J Gastrointestin Liver Dis*, **18**, 337-343.
- Probert, C. S. J., Jones, P. R. H. and Ratcliffe, N. M.** (2004). A novel method for rapidly diagnosing the causes of diarrhoea. *Gut*, **53**, 58-61.
- Putignani L, Menichella D.** (2010). Global distribution, public health and clinical impact of the protozoan pathogen *Cryptosporidium*. *Interdiscip Perspect Infect Dis*.
- Rahmouni, I., Essid, R., Aoun, K. and Bouratbine, A.** (2014). Glycoprotein 60 diversity in *Cryptosporidium parvum* causing human and cattle cryptosporidiosis in the rural region of Northern Tunisia. *Am J Trop Med Hyg*, **90**, 346-350.
- Ratanapo, S., Munghin, M., Soontrapa, S., Faithed, C., Siripattanapipong, S., Rangsin, R., Naaglor, T., Piyaraj, P., Taamasri, P. and Leelayoova, S.** (2008). Multiple modes of transmission of giardiasis in primary schoolchildren of a rural community, Thailand. *Am J Trop Med Hyg*, **78**, 611-615.

- Read, C. M., Monis, P. T. and Andrew Thompson, R. C.** (2004). Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infect Genet Evol*, **4**, 125-130.
- Roberts, L. and Zeibig, E.** (2013). Specimen Collection and Processing. In *Clinical Parasitology - a practical approach* (ed. Zeibig, E.), Elsevier, St Louis, Missouri.
- Robertson, L. J., Hanevik, K., Escobedo, A. A., Mørch, K. and Langeland, N.** (2010). Giardiasis- why do the symptoms sometimes never stop? *Trends Parasitol*, **26**, 75-82.
- Robertson, L. J., Hermansen, L., Gjerde, B. K., Strand, E., Alvsvåg, J. O. and Langeland, N.** (2006). Application of genotyping during an extensive outbreak of waterborne giardiasis in Bergen, Norway, during autumn and winter 2004. *Appl Environ Microbiol*, **72**, 2212-2217.
- Robertson LJ, Forberg T, Hermansen L, Gjerde BK, Langeland N. (2007). Molecular characterisation of *Giardia* isolates from clinical infections following a waterborne outbreak. *J Infect*, **55**, 79-88.
- Rutgeerts P, Van Assche G, Vermeire S, et al.** (April 2005). Ornidazole for prophylaxis of postoperative Crohn's disease recurrence: a randomized, double-blind, placebo-controlled trial. *Gastroenterol*, **128** (4), 856-61.
- Ryan, U., O'Hara, A. and Xiao, L.** (2004). Molecular and biological characterization of a *Cryptosporidium molnari*-like isolate from a guppy (*Poecilia reticulata*). *Appl Environ Microbiol*, **70**, 3761-3765.
- Ryan, Kenneth J.; Ray, C. George** (2004). Sherris Medical Microbiology: An Introduction to Infectious Disease (4th ed.). New York: McGraw-Hill. pp. 727-730.
- Ryan U, Cacciò SM.** (2013). Zoonotic potential of Giardia. *Int J Parasitol*, **43**, 943 -56.

- Sahagun, J., Clavel, A., Goni, P., Seral, C., Llorente, M. T., Castillo, F. J., Capilla, S., Arias, A. and Gomez-Lus, R.** (2008). Correlation between the presence of symptoms and the *Giardia duodenalis* genotype. *Eur J Clin Microbiol Infect Dis*, **27**, 81-83.
- Saitou N, Nei M.** (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, **4**, 406-25.
- Saksirisampant, W., Boontanom, P., Mungthin, M., Tan-Ariya, P., Lamchuan, D., Siripattanapipong, S. and Leelayoova, S.** (2012). Prevalence of giardiasis and genotypic characterization of *Giardia duodenalis* in hilltribe children, Northern Thailand. *Trop Biomed*, **29**, 331-338.
- Saksirisampant, W., Nuchprayoon, S., Wiwanitkit, V., Yenthakam, S. and Ampavasiri, A.** (2003). Intestinal parasitic infestations among children in an orphanage in Pathum Thani province. *J Med Assoc Thai*, **86 Suppl 2**, S263-270.
- Salyer, S. J., Gillespie, T. R., Rwego, I. B., Chapman, C. A. and Goldberg, T. L.** (2012). Epidemiology and molecular relationships of *Cryptosporidium* spp. in people, primates, and livestock from Western Uganda. *PLoS Negl Trop Dis*, **6**, e1597.
- Samokhin, A. and Revelsky, I.** (2011). Distinguishing by principal component analysis of o-xylene, m-xylene, p-xylene and ethylbenzene using electron ionization mass spectrometry. *Eur J Mass Spectrom (Chichester, Eng)*, **17**, 477-480.
- Samra, H. A., McGrath, J. M., Wehbe, M. and Clapper, J.** (2012). Epigenetics and family-centered developmental care for the preterm infant. *Adv Neonatal Care*, **12 (Suppl 5)**, S2-9.
- Savin, C., Sarfati, C., Menotti, J., Jaouhari, J., Wurtzer, S., Garin, Y. J. and Derouin, F.** (2008). Assessment of cryptodiag for diagnosis of cryptosporidiosis and genotyping *Cryptosporidium* species. *J Clin Microbiol*, **46**, 2590-2594.
- Savioli, L., Smith, H. and Thompson, A.** (2006). *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative'. *Trends Parasitol*, **22**, 203-208.

- Scorza AV, Ballweber LR, Tangtrongsup S, Panuska C, Lappin MR. (2012). Comparisons of mammalian *Giardia duodenalis* assemblages based on the bgirdin, glutamate dehydrogenase and triose phosphate isomerase genes. *Vet Parasitol*, **189**, 182–188.
- Shariff, M., Deb, M., Singh, R. and Singh, K. (2002). *Cryptosporidium* infection in children with diarrhoea of acute onset. *J Trop Pediatr*, **48**, 187-188.
- Shirley DA, Moonah SN, Kotloff KL. (2012).** Burden of disease from cryptosporidiosis. *Curr Opin Infect Dis*, **25**, 555-563
- Shukla, R., Giraldo, P., Kraliz, A., Finnigan, M. and Sanchez, A. L. (2006).** *Cryptosporidium* spp. and other zoonotic enteric parasites in a sample of domestic dogs and cats in the Niagara region of Ontario. *Can Vet J*, **47**, 1179-1184.
- Sinoun, M., R. Tsuyuoka, D. Socheat, A. Montresor, and K. Palmer. (2005).** Financial costs of deworming children in all primary schools in Cambodia. *Trans R Soc Trop Med Hyg*, **99**, 664-8.
- Siwila, J., Phiri, I. G., Enemark, H. L., Nchito, M. and Olsen, A. (2011).** Seasonal prevalence and incidence of *Cryptosporidium* spp. and *Giardia duodenalis* and associated diarrhoea in children attending pre-school in Kafue, Zambia. *Trans R Soc Trop Med Hyg*, **105**, 102-108.
- Slapeta, J. (2006).** *Cryptosporidium* species found in cattle: a proposal for a new species. *Trends Parasitol*, **22**, 469-474.
- Smith, H. V., Caccio, S. M., Tait, A., McLauchlin, J. and Thompson, R. C. (2006).** Tools for investigating the environmental transmission of *Cryptosporidium* and *Giardia* infections in humans. *Trends Parasitol*, **22**, 160-167.
- Smith, H. V., Campbell, B. M. and Grimason, A. M. (2003).** A microscopic system with a dual band filter for the simultaneous enumeration of *Cryptosporidium parvum* oocysts and sporozoites. *Water Res*, **37**, 2525-2527; author reply 2528-2530.

- Smith, H. V., Nichols, R. A., Mallon, M., Macleod, A., Tait, A., Reilly, W. J., Browning, L. M., Gray, D., Reid, S. W. and Wastling, J. M.** (2005). Natural *Cryptosporidium hominis* infections in Scottish cattle. *Vet Rec*, **156**, 710-711.
- Snelling, W. J., Xiao, L., Ortega-Pierres, G., Lowery, C. J., Moore, J. E., Rao, J. R., Smyth, S., Millar, B. C., Rooney, P. J., Matsuda, M., Kenny, F., Xu, J. and Dooley, J. S.** (2007). Cryptosporidiosis in developing countries. *J Infect Dev Ctries*, **1**, 242-256.
- Soave R, Johnson WD Jr.** (1988). *Cryptosporidium* and *Isospora belli* infections. *J Infect Dis*, **157**, 225.
- Soba, B., Petrovec, M., Mioc, V. and Logar, J.** (2006). Molecular characterisation of *Cryptosporidium* isolates from humans in Slovenia. *Clin Microbiol Infect*, **12**, 918-921.
- Sousa MC, Morais JB, Machado JE, Poiaraes-da-Silva J.** (2006). Genotyping of *Giardia lamblia* human isolates from Portugal by PCR-RFLP and sequencing. *J Eukaryot Microbiol*, **53** (Suppl 1), S174-6.
- Sprong, H., Cacciò, S.M. and van der Giessen, J. W.B.** (2009). Identification of zoonotic genotypes of *Giardia duodenalis*. *PLoS Negl Trop Dis*, **3**, 1-12.
- Stein S.E. and Scott D.R.** (1994). Optimisation and testing of mass spectral library search algorithms for compound identification. *J Am Soc Mass Spectrom*, **5**, 859-66.
- Stuart JM, Orr HJ, Warburton FG, Jeyakanth S, Pugh C, Morris I, Sarangi J, Nichols G.** (2003). Risk factors for sporadic giardiasis: a case-control study in southwestern England. *Emerg Infect Dis*, **9** (2), 229-233.
- Sulaiman, I. M., Fayer, R., Bern, C., Gilman, R. H., Trout, J. M., Schantz, P. M., Das, P., Lal, A. A. and Xiao, L.** (2003). *Triosephosphate isomerase* gene characterization and potential zoonotic transmission of *Giardia duodenalis*. *Emerg Infect Dis*, **9**, 1444-1452.

- Sulaiman, I. M., Hira, P. R., Zhou, L., Al-Ali, F. M., Al-Shelahi, F. A., Shweiki, H. M., Iqbal, J., Khalid, N. and Xiao, L.** (2005). Unique endemicity of cryptosporidiosis in children in Kuwait. *J Clin Microbiol*, **43**, 2805-2809.
- Tamer, G. S., Erdogan, S. and Willke, A.** (2008). The frequency of the presence of intestinal parasites in students of Arslanbey Primary School. *Turkiye Parazitol Derg*, **32**, 130-133.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S.** (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*, **30**, 2725-2729.
- Tanriverdi, S., Grinberg, A., Chalmers, R. M., Hunter, P. R., Petrovic, Z., Akiyoshi, D. E., London, E., Zhang, L., Tzipori, S., Tumwine, J. K. and Widmer, G.** (2008). Inferences about the global population structures of *Cryptosporidium parvum* and *Cryptosporidium hominis*. *Appl Environ Microbiol*, **74**, 7227-7234.
- Teodorovic S, Braverman JM, Elmendorf HG.** (2007). Unusually low levels of genetic variation among *Giardia lamblia* isolates. *Eukaryot Cell*, **6**, 1421-1430.
- Thompson, R. C.** (2004). The zoonotic significance and molecular epidemiology of *Giardia* and giardiasis. *Vet Parasitol*, **126**, 15-35.
- Thompson, R. C., Hopkins, R. M. and Homan, W. L.** (2000). Nomenclature and genetic groupings of *Giardia* infecting mammals. *Parasitol Today*, **16**, 210-213.
- Thompson, R. C. and Monis, P. T.** (2004a). Variation in *Giardia*: implications for taxonomy and epidemiology. *Adv Parasitol*, **58**, 69-137.
- Thompson, R. C. A., O'Handley, R. and Palmer, C. S.** (2008). The public health and clinical significance of *Giardia* and *Cryptosporidium* in domestic animals. *Veterinary journal*, **177**, 18-25.
- Thompson RC, Smith A.** (2011). Zoonotic enteric protozoa. *Vet Parasitol*, **182**, 70-8.

- Tiangtip, R. and Jongwutiwes, S.** (2002). Molecular analysis of *Cryptosporidium* species isolated from HIV-infected patients in Thailand. *Trop Med Int Health*, **7**, 357-364.
- Traub, R. J., Monis, P. T., Robertson, I., Irwin, P., Mencke, N. and Thompson, R. C.** (2004). Epidemiological and molecular evidence supports the zoonotic transmission of *Giardia* among humans and dogs living in the same community. *Parasitol*, **128**, 253-262.
- Trotz-Williams, L. A., Martin, D. S., Gatei, W., Cama, V., Peregrine, A. S., Martin, S. W., Nydam, D. V., Jamieson, F. and Xiao, L.** (2006). Genotype and subtype analyses of *Cryptosporidium* isolates from dairy calves and humans in Ontario. *Parasitol Res*, **99**, 346-352.
- Trout, J. M., Santin, M. and Fayer, R.** (2003). Identification of assemblage A *Giardia* in white-tailed deer. *J Parasitol*, **89**, 1254-1255.
- Trout JM, Santin M, Greiner E, Fayer R.** (2004). Prevalence of *Giardia duodenalis* genotypes in pre-weaned dairy calves. *Vet Parasitol* **124**, 179-186.
- Tumwine, J. K., Kekitiinwa, A., Nabukeera, N., Akiyoshi, D. E., Rich, S. M., Widmer, G., Feng, X. and Tzipori, S.** (2003). *Cryptosporidium parvum* in children with diarrhea in Mulago Hospital, Kampala, Uganda. *Am J Trop Med Hyg*, **68**, 710-715.
- Tungtrongchitr, A., Sookrung, N., Indrawattana, N., Kwangsi, S., Ongrotchanakun, J. and Chaicumpa, W.** (2010). *Giardia intestinalis* in Thailand: identification of genotypes. *J Health Popul Nutr*, **28**, 42-52.
- UNICEF.** UNICEF Statistics for Cambodia. http://www.unicef.org/infobycountry/cambodia_statistics.html- 73 (23 September 2011, date last accessed).
- UNICEF** (2009), posting date. UNICEF Statistics for Cambodia.
- United State Environment Protection Agency.** Emergency Disinfection of Drinking Water. Retrieved 21 June 2015. Retrieved 24 February 2011.

- Upcroft, P. & Upcroft, J.A.** (2001). Drug targets and mechanisms of resistance in anaerobic protozoa. *Clin Microbiol Rev*, **14**, 150-164.
- Urbani, C., Sinoun, M., Socheat, D., Pholsena, K., Strandgaard, H., Odermatt, P. and Hatz, C.** (2002). Epidemiology and control of *mekongi* schistosomiasis. *Acta Trop*, **82**, 157-168.
- van Keulen H, Macechko PT, Wade S, Schaaf S, Wallis PM, Erlandsen SL.** (2002). Presence of human *Giardia* in domestic, farm and wild animals, and environmental samples suggests a zoonotic potential for giardiasis. *Vet Parasitol*, 108, 97-107.
- Vas, G. and Vekey, K.** (2004). Solid-phase microextraction: a powerful sample preparation tool prior to mass spectrometric analysis. *J Mass Spectrom*, **39**, 233-254.
- Vargas M, Gascon J, Casals C, Schellenberg D, Kahigwa E, et al.** (2004). Etiology of diarrhea in children less than five years of age in Ifakara, Tanzania. *Am J Trop Med Hyg*, **70**, 536-539.
- Verweij, J. J.** (2014). Application of PCR-based methods for diagnosis of intestinal parasitic infections in the clinical laboratory. *Parasitol*, **29**, 1-10
- Verweij, J. J., Schinkel, J., Laeijendecker, D., van Rooyen, M. A., van Lieshout, L. and Polderman, A. M.** (2003). Real-time PCR for the detection of *Giardia lamblia*. *Mol Cell Probes*, **17**, 223-225.
- Volotao, A. C., Costa-Macedo, L. M., Haddad, F. S., Brandao, A., Peralta, J. M. and Fernandes, O.** (2007). Genotyping of *Giardia duodenalis* from human and animal samples from Brazil using *beta-giardin* gene: a phylogenetic analysis. *Acta Trop*, **102**, 10-19.
- Water Research Foundation (formerly AwwaRF).** "Ultraviolet Disinfection and Treatment" Archived from the original on 24 January 2009.
- Weerapol, T.** (2010). Molecular epidemiology of *Giardia duodenalis*. Vol. PhD, Veterinary Science University of Liverpool, Liverpool.

- Wegayehu, T., Adamu, H. and Petros, B.** (2013a). Prevalence of *Giardia duodenalis* and *Cryptosporidium* species infections among children and cattle in North Shewa Zone, Ethiopia. *BMC Infect Dis*, **13**, 419.
- Wegayehu, T., Tsalla, T., Seifu, B. and Teklu, T.** (2013b). Prevalence of intestinal parasitic infections among highland and lowland dwellers in Gamo area, South Ethiopia. *BMC Public Health*, **13**, 151.
- Wellington O, Chika O, Teslim O, Oladipo O, Adetayo F, Godswill I.** (2008). *Cryptosporidium* and other Intestinal Protozoans in Children with Diarrhoea in Lagos, Nigeria. *Internet J Trop Med*, **5**(2).
- WHO Model List of Essential Medicines** (19th List) (PDF). World Health Organization. April 2015. Retrieved 8 December 2016.
- Wielinga, P. R., de Vries, A., van der Goot, T. H., Mank, T., Mars, M. H., Kortbeek, L. M. and van der Giessen, J. W.** (2008). Molecular epidemiology of *Cryptosporidium* in humans and cattle in The Netherlands. *Int J Parasitol*, **38**, 809-817.
- Wielinga CM, Thompson RC.** (2007). Comparative evaluation of *Giardia duodenalis* sequence data. *Parasitol*. **134**, 1795-1821.
- Wolfe, M. S.** (1992). Giardiasis. *Clin Microbiol Rev*, **5**, 93-100.
- Wongstitwilairoong, B., Srijan, A., Serichantalergs, O., Fukuda, C. D., McDaniel, P., Bodhidatta, L. and Mason, C. J.** (2007). Intestinal parasitic infections among pre-school children in Sangkhlaburi, Thailand. *Am J Trop Med Hyg*, **76**, 345-350.
- Workowski KA, Bolan GA.** (2015). Centers for Disease Control and Prevention (CDC). Sexually transmitted diseases treatment guidelines, **64**(33), 924.
- Wu Z. et. al.** (2003). Intraspecies polymorphism of *Cryptosporidium parvum* revealed by PCR-restriction fragment length polymorphism (RFLP) and RFLP-single-strand conformational polymorphism analyses. *Appl. Environ. Microbiol*, **69**, 4720-4726.

- WHO** <http://www.who.int/whr/en/> <http://www.who.int/whr/en/> (2014). World Health Report – Make Every Child and Mother Count. Facts and Figures from World Health Report 2005. Accessed April, 2014.
- WHO.** 2008. Review of the Epidemiological Profile of Helminthiases and their Control in the Western Pacific Region, 1997-2008.
- Xia, J., Psychogios, N., Young, N. and Wishart, D. S.** (2009). MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res*, **37**, 652-660.
- Xia, J. and Wishart, D. S.** (2011a). Metabolomic data processing, analysis, and interpretation using MetaboAnalyst. *Curr Protoc Bioinformatics*, **Chapter 14**, Unit 10.
- Xia, J. and Wishart, D. S.** (2011b). Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat Protoc*, **6**, 743-760.
- Xiao, L., Alderisio, K. A. and Jiang, J.** (2006a). Detection of *Cryptosporidium* oocysts in water: effect of the number of samples and analytic replicates on test results. *Appl Environ Microbiol*, **72**, 5942-5947.
- Xiao, L., Bern, C., Limor, J., Sulaiman, I., Roberts, J., Checkley, W., Cabrera, L., Gilman, R. H. and Lal, A. A.** (2001a). Identification of 5 types of *Cryptosporidium* parasites in children in Lima, Peru. *J Infect Dis*, **183**, 492-497.
- Xiao, L. and Fayer, R.** (2008). Molecular characterisation of species and genotypes of *Cryptosporidium* and *Giardia* and assessment of zoonotic transmission. *Int J Parasitol*, **38**, 1239-1255.
- Xiao L1, Fayer R, Ryan U, Upton SJ.** (2004). *Cryptosporidium* taxonomy: recent advances and implications for public health. *Clin Microbiol Rev*, **17**(1), 72-97.
- Xiao L, Feng Y.** (2008). Zoonotic cryptosporidiosis. *FEMS Immunol Med Microbiol*, **52**, 309-323.

- Xiao, L., Lal, A. A. and Jiang, J.** (2004). Detection and differentiation of *Cryptosporidium* oocysts in water by PCR-RFLP. *Methods Mol Biol*, **268**, 163-176.
- Xiao, L., Limor, J., Bern, C., Lal, A. A. and Epidemic Working, G.** (2001b). Tracking *Cryptosporidium parvum* by sequence analysis of small double-stranded RNA. *Emerg Infect Dis*, **7**, 141-145.
- Xiao, L., Moore, J. E., Ukoh, U., Gatei, W., Lowery, C. J., Murphy, T. M., Dooley, J. S., Millar, B. C., Rooney, P. J. and Rao, J. R.** (2006b). Prevalence and identity of *Cryptosporidium* spp. in pig slurry. *Appl Environ Microbiol*, **72**, 4461-4463.
- Xiao, L., Ryan, U. M.** (2008). Molecular epidemiology. In: Fayer, R., Xiao, L. (Eds.), *Cryptosporidium* and Cryptosporidiosis. CRC Press and IWA Publishing, Boca Raton, FL, 119-171.
- Yong TS, Park SJ, Hwang UW, Yang HW, Lee KW, Min DY, et al.** (2000). Genotyping of *Giardia lamblia* isolates from humans in China and Korea using ribosomal DNA Sequences. *J Parasitol*, **86**, 887-91.
- Yang R, Lee J, Ng J, Ryan U.** (2010). High prevalence *Giardia duodenalis* assemblage B and potentially zoonotic sub-assemblages in sporadic human cases in Western Australia. *Int J Parasitol*, **40**, 293-7.
- Zeynudin, A., Hemalatha, K. and Kannan, S.** (2013). Prevalence of opportunistic intestinal parasitic infection among HIV infected patients who are taking antiretroviral treatment at Jimma Health Center, Jimma, Ethiopia. *Eur Rev Med Pharmacol Sci*, **17**, 513-516.
- Zhou, L., Fayer, R., Trout, J. M., Ryan, U. M., Schaefer, F. W., 3rd and Xiao, L.** (2004). Genotypes of *Cryptosporidium* species infecting fur-bearing mammals differ from those of species infecting humans. *Appl Environ Microbiol*, **70**, 7574-7577.

Zintl, A., Proctor, A. F., Read, C., Dewaal, T., Shanaghy, N., Fanning, S. and Mulcahy, G. (2009).
The prevalence of *Cryptosporidium* species and subtypes in human faecal samples in Ireland.
Epidemiol Infect, **137**, 270-277.